The Annals of Mathematical Statistics
1971, Vol. 42, No. 1, 315–324

CONVERGENCE OF CONDITIONAL EXPECTATIONS

BY P. GÄNSSLER AND J. PFANZAGL

University of Cologne

0. Introduction. Let $(X, \mathscr{F})$ be a measurable space and $P_n|\mathscr{F}$, $n = 0, 1, 2, \cdots$, a family of probability measures such that $(P_n)_{n \in \mathbb{N}}$ converges to P_0 in some appropriate sense. Let $\mathscr{F}_0 \subset \mathscr{F}$ be an arbitrary sub-σ-field. For $f \in \bigcap_{n=0}^{\infty} \mathscr{L}_1(X, \mathscr{F}, P_n)$ let $p_n(f, \cdot)$ denote a conditional expectation of f relative to P_n, given $\mathscr{F}_0$. Our problem is to give sufficient conditions under which $(p_n(f, \cdot))_{n \in \mathbb{N}}$ converges to $p_0(f, \cdot)$ in some sense.

1. The results. Though the convergence of conditional expectations has been treated by a number of authors (see references) and various sufficient conditions have been given, the following rather natural condition seems to have been overlooked.

Let $\mu|\mathscr{F}$ be a σ-finite measure dominating $P_n|\mathscr{F}$ for all $n = 0, 1, 2, \cdots$. Let h_n be a density of $P_n|\mathscr{F}$ with respect to $\mu|\mathscr{F}$ and h_{0n} a density of $P_n|\mathscr{F}_0$ with respect to $\mu|\mathscr{F}_0$.

THEOREM 1. *Assume that*
(i) $(h_n)_{n \in \mathbb{N}}$ *converges to* h_0 μ-a.e.
(ii) $(h_{0n})_{n \in \mathbb{N}}$ *converges to* h_{00} μ-a.e.

Then

(1) *For abitrary versions of the conditional expectations*: $(p_n(f, \cdot))_{n \in \mathbb{N}}$ *converges to* $p_0(f, \cdot)$ P_0-a.e. *for any* $\mathscr{F}$*-measurable, bounded function* f.

(2) *If a regular conditional probability relative to* μ, *given* $\mathscr{F}_0$, *exists, then there exists a regular conditional probability* $p_n^*|\mathscr{F} \times X$ *relative to* P_n, *given* $\mathscr{F}_0$, *such that* $(\sup_{A \in \mathscr{F}} |p_n^*(A, \cdot) - p_0^*(A, \cdot)|)_{n \in \mathbb{N}}$ *converges to* 0 P_0-a.e.

(3) *If, in addition,* $\mathscr{F}$ *is countably generated, the uniform convergence asserted in* (2) *holds for all regular conditional probabilities* $p_n|\mathscr{F} \times X$ *relative to* P_n, *given* $\mathscr{F}_0$.

We remark that the conditions on the densities depend neither on the particular dominating measure μ nor on the particular versions chosen. Hence w.l.g. $\mu(X) = 1$.

PROOF. (i) Let $B_n = \{x \in X: h_{0n}(x) > 0\}$ and $q_n(x) = ((h_n(x)/h_{0n}(x))1_{B_n}(x) + 1_{\bar{B}_n}(x)$, $n = 0, 1, 2, \cdots$, (where $\bar{B}_n$ denotes the complementary set of B_n in X). Since $P_n(\bar{B}_n) = 0$, we have $h_n(x) = 0$ for μ-a.a. $x \in \bar{B}_n$ and therefore $h_{0n} q_n = h_n$ μ-a.e. For $g \in \mathscr{L}_1(X, \mathscr{F}, \mu)$ let $\mu(g, \cdot) \in \mathscr{L}_1(X, \mathscr{F}_0, \mu)$ denote a conditional expectation of g relative to μ, given $\mathscr{F}_0$. As $\mu(h_n, \cdot)$ is a density of $P_n|\mathscr{F}_0$ with respect to $\mu|\mathscr{F}_0$, we have $h_{0n} = \mu(h_n, \cdot)$ μ-a.e. and therefore $\mu(q_n, \cdot) = 1$ μ-a.e. This implies in particular that $\mu(q_n) = 1$ so that all functions $q_n 1_B$ are μ-integrable. (Here and in the following, $\mu(g)$ means $\int g(\xi)\, \mu(d\xi)$.)

Received September 23, 1969.

We shall show that $\mu(q_n f, \cdot)$ is a conditional expectation of f relative to P_n, given $\mathscr{F}_0$. As $\mu(q_n f, \cdot)$ is $\mathscr{F}_0$-measurable by definition, this follows immediately from

$$\begin{aligned} P_n(\mu(q_n f, \cdot)1_A) &= \mu(h_{0n}\mu(q_n f, \cdot)1_A) \\ &= \mu(\mu(h_{0n} q_n f, \cdot)1_A) = \mu(\mu(h_n f, \cdot)1_A) \\ &= \mu(h_n f 1_A) = P_n(f 1_A) \qquad \text{for every} \quad A \in \mathscr{F}_0. \end{aligned}$$

(ii) Now we shall show that $(q_n 1_{B_0})_{n\in\mathbb{N}} \to q_0 1_{B_0}$ μ-a.e. Let N denote the μ-null set comprising all those $x \in X$ for which $(h_n(x))_{n\in\mathbb{N}} \to h_0(x)$ or $(h_{0n}(x))_{n\in\mathbb{N}} \to h_{00}(x)$ fails to be true. $x \in \bar{N} \cap B_0$ implies $h_{0n}(x) > 0$ and therefore $q_n(x) = h_n(x)/h_{0n}(x)$ for all sufficiently large n, whence $\lim_{n\in\mathbb{N}} q_n(x) 1_{B_0}(x) = h_0(x)/h_{00}(x) = q_0(x) 1_{B_0}(x)$. If $x \in \bar{B}_0$, $\lim_{n\in\mathbb{N}} q_n(x) 1_{B_0}(x) = 0 = q_0(x) 1_{B_0}(x)$.

(iii) Now we shall show that $(\mu(|q_n - q_0|, \cdot))_{n\in\mathbb{N}}$ converges to 0 P_0-a.e. The proof is similar to the proof of Scheffé's Lemma (see Lehmann, page 352).

We have $0 \leqq (q_n 1_{B_0} - q_0 1_{B_0})^- \leqq q_0 1_{B_0}$. Hence $((q_n 1_{B_0} - q_0 1_{B_0})^-)_{n\in\mathbb{N}} \to 0$ μ-a.e. implies (cf. Doob, pages 23–24 CE_5)

$$(4) \qquad (\mu(q_n 1_{B_0} - q_0 1_{B_0})^-, \cdot))_{n\in\mathbb{N}} \to 0 \ \ \mu\text{-a.e.}$$

As $(q_n 1_{B_0} - q_0 1_{B_0})^+ = (q_n 1_{B_0} - q_0 1_{B_0}) + (q_n 1_{B_0} - q_0 1_{B_0})^-$, we have $\mu((q_n 1_{B_0} - q_0 1_{B_0})^+, \cdot) = \mu((q_n 1_{B_0} - q_0 1_{B_0}), \cdot) + \mu((q_n 1_{B_0} - q_0 1_{B_0})^-), \cdot)$ μ-a.e. Furthermore, for μ-a.a. $x \in X$: $\mu((q_n 1_{B_0} - q_0 1_{B_0}), x) = \mu(q_n 1_{B_0}, x) - \mu(q_0 1_{B_0}, x) = 0$ (since $\mu(q_n, x) = 1$ for μ-a.a. $x \in X$ and all $n = 0, 1, 2, \cdots$). Hence (4) implies

$$(5) \qquad (\mu((q_n 1_{B_0} - q_0 1_{B_0})^+, \cdot)_{n\in\mathbb{N}} \to 0 \ \ \mu\text{-a.e.}$$

(4) and (5) together imply

$$(\mu(|q_n 1_{B_0} - q_0 1_{B_0}|, \cdot)_{n\in\mathbb{N}} \to 0 \ \ \mu\text{-a.e.}$$

As $\mu(|q_n 1_{B_0} - q_0 1_{B_0}|, \cdot) = \mu(|q_n - q_0|, \cdot) 1_{B_0}$ μ-a.e., this implies

$$(6) \qquad (\mu(|q_n - q_0|, \cdot))_{n\in\mathbb{N}} \to 0 \ \ P_0\text{-a.e.}$$

(iv) Let f be an $\mathscr{F}$-measurable, bounded function. Then

$$\begin{aligned} |\mu(q_n f, \cdot) - \mu(q_0 f, \cdot)| &\leqq \mu(|q_n f - q_0 f|, \cdot) \\ &\leqq \mu(|q_n - q_0|, \cdot) \sup_{x\in X} |f(x)| \ \ \mu\text{-a.e.} \end{aligned}$$

implies $\lim_{n\in\mathbb{N}} \mu(q_n f, \cdot) = \mu(q_0 f, \cdot)$ P_0-a.e. This proves (1) for the particular version $p_n'(f, \cdot) = \mu(q_n f, \cdot)$.

(v) To see that (1) holds for arbitrary versions, we proceed as follows: Let $p_n(f, \cdot)$, $n = 0, 1, 2, \cdots$, be arbitrary versions of the conditional expectations of f relative to P_n, given $\mathscr{F}_0$. Let $A_n = \{x \in X: p_n'(f, x) \neq p_n(f, x)\}$, $n = 0, 1, 2, \cdots$. We have $A_n \in \mathscr{F}_0$ and $P_n(A_n) = 0$, $n = 0, 1, 2, \cdots$, whence $P_0(\limsup_{n\in\mathbb{N}} (A_n \cup A_0)) = 0$ by Lemma 1 (see below). Hence for P_0-a.a. $x \in X$ there exists $n(x) \in \mathbb{N}$ such that $p_0(f, x) = p_0'(f, x)$ and $p_n(f, x) = p_n'(f, x)$ for all $n \geqq n(x)$. This implies (1).

(vi) If $\tilde{\mu}\,|\,\mathscr{F}\times X$ is a regular conditional probability relative to μ, given $\mathscr{F}_0$, and if $\mu(g,\cdot)=\int g(\xi)\,\tilde{\mu}(d\xi,\cdot)$, then

$$|\mu(q_n 1_A,\cdot)-\mu(q_0 1_A,\cdot)| \leqq \mu(|q_n 1_A - q_0 1_A|,\cdot) \leqq \mu(|q_n - q_0|,\cdot)$$

holds everywhere, whence

$$\sup_{A\in\mathscr{F}} |\mu(q_n 1_A,\cdot)-\mu(q_0 1_A,\cdot)| \leqq \mu(|q_n - q_0|,\cdot).$$

As for μ-a.a. $x\in X$ $\mu(q_n, x)=1$, $n=0,1,2,\cdots$ and as μ dominates P_0, there is a P_0-null set N such that $\mu(q_n, x)=1$ for $x\in\bar{N}$, $n=0,1,2,\cdots$. Then (6) implies (2) for

$$\begin{aligned} p_n^*(A,x) &= \mu(q_n 1_A, x) \qquad &\text{if}\quad x\in\bar{N}\\ &= P_n(A) \qquad &\text{if}\quad x\in N. \end{aligned}$$

(vii) For $n=0,1,2,\cdots$ let $p_n\,|\,\mathscr{F}\times X$ be an arbitrary regular version of the conditional probability relative to P_n, given $\mathscr{F}_0$. If $\mathscr{F}$ is countably generated, there exists a countable algebra, say $\mathscr{H}$, generating $\mathscr{F}$.

For $n=0,1,2,\cdots$ let

$$M_n = \bigcup_{A\in\mathscr{H}} \{x\in X: p_n(A,x)\neq p_n^*(A,x)\}.$$

We have $M_n\in\mathscr{F}_0$ and $P_n(M_n)=0$ for $n=0,1,2,\cdots$. According to Lemma 1 this implies $P_0(\limsup_{n\in\mathbb{N}}(M_n\cup M_0))=0$. As

$$\begin{aligned} &\{x\in X: \sup_{A\in\mathscr{H}} |p_n(A,x)-p_0(A,x)| \neq \sup_{A\in\mathscr{H}} |p_n^*(A,x)-p_0^*(A,x)|\}\\ &\qquad \subset \bigcup_{A\in\mathscr{H}} \{x\in X: |p_n(A,x)-p_0(A,x)| \neq |p_n(A,x)-p_0^*(A,x)|\} \subset M_n\cup M_0, \end{aligned}$$

the relation $x\notin\limsup_{n\in\mathbb{N}}(M_n\cup M_0)$ implies for all $n\geqq n(x)$:

$$\sup_{A\in\mathscr{H}} |p_n(A,x)-p_0(A,x)| = \sup_{A\in\mathscr{H}} |p_n^*(A,x)-p_0^*(A,x)|. \tag{7}$$

As p_n and p_n^* are regular, the suprema over $\mathscr{H}$ equal the suprema over $\mathscr{F}$. (This follows easily from the approximation theorem. See in particular Halmos, page 58, Problem 8.) Hence (7) implies

$$\sup_{A\in\mathscr{F}} |p_n(A,x)-p_0(A,x)| = \sup_{A\in\mathscr{F}} |p_n^*(A,x)-p_0^*(A,x)|$$

$$\text{for all}\quad x\notin\limsup_{n\in\mathbb{N}}(M_n\cup M_0) \quad\text{and all}\quad n\geqq n(x).$$

Together with (2) this implies

$$\lim_{n\in\mathbb{N}} \sup_{A\in\mathscr{F}} |p_n(A,x)-p_0(A,x)| = 0 \qquad \text{for}\quad P_0\text{-a.a. } x\in X.$$

REMARK. If $p_n^*\,|\,\mathscr{F}\times X$, $n=0,1,2,\cdots$, are regular versions of the conditional probability relative to P_n, given $\mathscr{F}_0$, then $p_n(f,x)=\int f(\xi)p_n^*(d\xi,x)$ is a conditional expectation of f relative to P_n, given $\mathscr{F}_0$ (see Doob, page 27, Theorem 9.1). If (2) of Theorem 1 holds for p_n^*, then

$$\lim_{n\in\mathbb{N}} \sup_{f\in\mathfrak{F}} |p_n(f,x)-p_0(f,x)| = 0 \qquad \text{for}\quad P_0\text{-a.a. } x\in X$$

for any uniformly bounded set $\mathfrak{F}$ of $\mathscr{F}$-measurable functions. This follows immediately from

$$\sup_{f\in\mathfrak{F}}|p_n(f,x)-p_0(f,x)| \leqq \sup_{A\in\mathscr{F}}|p_n{}^*(A,x)-p_0{}^*(A,x)|.$$

The same remark holds for Theorem 2 and Theorem 3.

The following Example 1 shows that (cf. (vi) in the proof of Theorem 1) it is not true in general that for all regular conditional probabilities $\tilde{\mu}\,|\,\mathscr{F}\times X$ relative to μ, given $\mathscr{F}_0$ and all densities h of some probability measure $P\,|\,\mathscr{F}$ with respect to $\mu\,|\,\mathscr{F}$

$$\int q(y)\,\tilde{\mu}(dy,x) = 1 \qquad \text{for all} \quad x\in X,$$

where
$$q(x) = \frac{h(x)}{h_0(x)}1_{\{y\in X:\, h_0(y)>0\}}(x)+1_{\{y\in X:\, h_0(y)\leqq 0\}}(x)$$

and
$$h_0(x) = \int h(y)\,\tilde{\mu}(dy,x).$$

EXAMPLE 1. Let $X = [0, 1]$, $\mathscr{F}$ the Borel sets in $[0, 1]$ and $\mu\,|\,\mathscr{F}$ the Lebesgue measure in $[0, 1]$. Then

$$\begin{aligned} \tilde{\mu}(A,x) &= 1_A(x) && \text{if} \quad x\in(0,1], \\ &= \tfrac{1}{2}1_A(0)+\tfrac{1}{2}\mu(A) && \text{if} \quad x=0 \end{aligned}$$

defines a regular conditional probability $\tilde{\mu}\,|\,\mathscr{F}\times X$ of μ, given $\mathscr{F}_0 = \mathscr{F}$.

Let $P\,|\,\mathscr{F}$ be a probability measure such that there exists a density h of $P\,|\,\mathscr{F}$ with respect to $\mu\,|\,\mathscr{F}$ with $h(x) > 0$ for $x\in(0, 1]$ and $h(0) = 0$. We have $\mu(h) = 1$. Hence

$$\begin{aligned} h_0(x) = \int h(y)\,\tilde{\mu}(dy,x) &= h(x) && \text{if} \quad x\in(0,1]; \\ &= \tfrac{1}{2} && \text{if} \quad x=0. \end{aligned}$$

Then

$$\begin{aligned} q(x) = h(x)/h_0(x) &= 1 && \text{if} \quad x\in(0,1]; \\ &= 0 && \text{if} \quad x=0. \end{aligned}$$

We have

$$\begin{aligned} \int q(y)\,\tilde{\mu}(dy,x) &= q(x) = 1 && \text{if} \quad x\in(0,1]; \\ &= \tfrac{1}{2}q(0)+\tfrac{1}{2}\mu(q) = \tfrac{1}{2} && \text{if} \quad x=0. \end{aligned}$$

THEOREM 2. *Assume that*

(i) $(h_n)_{n\in\mathbb{N}}$ *converges to* h_0 μ-a.e.

(ii) $\mu(\sup_{n\in\mathbb{N}} h_n) < \infty$.

Then assertions (1)–(3) *of Theorem* 1 *hold.*

We remark that the condition on the densities depend neither on the particular dominating measure μ nor on the particular versions chosen. Hence w.l.g. $\mu(X) = 1$.

PROOF. For $g\in\mathscr{L}_1(X, \mathscr{F}, \mu)$ let $\mu(g, \cdot)\in\mathscr{L}_1(X, \mathscr{F}_0, \mu)$ denote a conditional expectation of g relative to μ, given $\mathscr{F}_0$. Condition (i) together with (ii) implies (cf. Doob, pages 23–24, CE_5)

$$(\mu(h_n, \cdot))_{n\in\mathbb{N}} \to \mu(h_0, \cdot) \quad \mu\text{-a.e.}$$

As $\mu(h_n, \cdot)$ is a density of $P_n | \mathscr{F}_0$ with respect to $\mu | \mathscr{F}_0$, Condition (ii) of Theorem 1 is fulfilled. Hence the assertion follows from Theorem 1.

For the case of parametrized families of probability measures we immediately obtain from Theorem 2 the following

COROLLARY 1. *Let $P_\vartheta | \mathscr{F}$, $\vartheta \in \Theta$, be a family of probability measures which is dominated by a σ-finite measure $\mu | \mathscr{F}$. Assume that Θ is endowed with a topology and that for each $\vartheta \in \Theta$ there exists a density h_ϑ of $P_\vartheta | \mathscr{F}$ with respect to $\mu | \mathscr{F}$ such that for each $\vartheta_0 \in \Theta$:*

(i) *$\vartheta \to h_\vartheta(x)$ is continuous in ϑ_0 for μ-almost all $x \in X$ (with the exceptional null-set possibly depending on ϑ_0).*

(ii) *There exist a neighborhood U_0 of ϑ_0 and $g_{\vartheta_0} \in \mathscr{L}_1(X, \mathscr{F}, \mu)$ such that for all $\vartheta \in U_0$, $h_\vartheta(x) \leqq g_{\vartheta_0}(x)$ for μ-almost all $x \in X$.*

Let $\mathscr{F}_0 \subset \mathscr{F}$ be an arbitrary sub-σ-field. Then $(\vartheta_n)_{n \in \mathbb{N}} \to \vartheta_0$ implies that

(8) *For arbitrary versions of the conditional expectations: $(p_{\vartheta_n}(f, \cdot))_{n \in \mathbb{N}}$ converges to $p_{\vartheta_0}(f, \cdot)$ P_{ϑ_0}-a.e. for any $\mathscr{F}$-measurable, bounded function f.*

(9) *If a regular conditional probability relative to μ, given $\mathscr{F}_0$, exists, then there exists a regular conditional probability $p^*_{\vartheta_n} | \mathscr{F} \times X$ relative to P_{ϑ_n}, given $\mathscr{F}_0$, such that $(\sup_{A \in \mathscr{F}} |p^*_{\vartheta_n}(A, x) - p_{\vartheta_0}(A, x)|)_{n \in \mathbb{N}}$ converges to 0 P_{ϑ_0}-a.e.*

(10) *If, in addition, $\mathscr{F}$ is countably generated, the uniform convergence asserted in (9) holds for all regular conditional probabilities $p_{\vartheta_n} | \mathscr{F} \times X$ relative to P_{ϑ_n}, given $\mathscr{F}_0$.*

The following Example 2 shows that the boundedness of f assumed in (1) of Theorem 1 and Theorem 2 is essential.

(1) can, however, be extended to cover unbounded functions f, if Condition (ii) in Theorem 2 is strengthened: By the same techniques it can be shown that $\sup_{n \in \mathbb{N}} h_n \in \mathscr{L}_p(X, \mathscr{F}, \mu)$ implies P_0-a.e. convergence of $(p_n(f, \cdot))_{n \in \mathbb{N}}$ to $p_0(f, \cdot)$ for all $f \in \mathscr{L}_1(X, \mathscr{F}, P_0) \cap \mathscr{L}_q(X, \mathscr{F}, \mu)$ $(1/p + 1/q = 1)$, because then $\sup_{n \in \mathbb{N}} (h_n f) \in \mathscr{L}_1(X, \mathscr{F}, \mu)$ so that (i) implies $(\mu(h_n f, \cdot))_{n \in \mathbb{N}} \to \mu(h_0 f, \cdot)$ μ-a.e. Together with $(\mu(h_n, \cdot))_{n \in \mathbb{N}} \to \mu(h_0, \cdot)$ μ-a.e. this implies the assertion by Lemma 2.

EXAMPLE 2. Let $X = \{0, 1, 2, \cdots\}$, $\mathscr{F} = \mathscr{P}(X)$ the power set, $\mathscr{F}_0 = \{\emptyset, X\}$. Let $P_n | \mathscr{F}$ be defined by $P_n\{0\} = 1 - 2^{-2n}$ and $P_n\{n\} = 2^{-2n}$ for $n \in \mathbb{N}$ and $P_0 | \mathscr{F}$ by $P_0\{0\} = 1$. Let $\mu | \mathscr{F}$ be defined by $\mu\{0\} = \frac{1}{2}$ and $\mu\{n\} = 2^{-(n+1)}$ for $n \in \mathbb{N}$. Then

$$\begin{aligned} h_n(m) &= 2(1 - 2^{-2n}) && \text{if } m = 0, \\ &= 2^{-(n-1)} && \text{if } m = n, \\ &= 0 && \text{otherwise}; \end{aligned}$$

$$\begin{aligned} h_0(m) &= 2 && \text{if } m = 0, \\ &= 0 && \text{otherwise}; \end{aligned}$$

are the corresponding densities of $P_n | \mathscr{F}$ with respect to $\mu | \mathscr{F}$, $n = 0, 1, 2, \cdots$.

We have $(h_n)_{n\in\mathbb{N}} \to h_0$ μ-a.e. Condition (ii) of Theorem 1 holds trivially (with $h_{0n} \equiv 1$ for all $n = 0, 1, 2, \cdots$) and even condition (ii) of Theorem 2 is fulfilled: $\mu(\sup_{n\in\mathbb{N}} h_n) = \sum_{m=0}^{\infty} 2^{-2m} < \infty$. However $p_n(f_0, \cdot) = P_n(f_0) = 1 \neq 0 = P_0(f_0) = p_0(f_0, \cdot)$ for all $n \in \mathbb{N}$, where f_0 is the function defined by $f_0(0) = 0$, $f_0(n) = 2^{2n}$, $n \in \mathbb{N}$.

We remark that Theorem 1 (1) generalizes Theorem 4.4 of Trumbo which asserts a.e. convergence for particular versions of the conditional expectations for the particular case of X being a Cartesean product and $\mathscr{F}$ being the product of $\mathscr{F}_0$ and another σ-field.

Example 4.1 in Trumbo shows that condition (i) of Theorem 1 does not imply condition (ii) and that, furthermore, condition (i) alone is not sufficient to ensure even the existence of a.e. convergent versions.

The other question is whether, retaining condition (ii) of Theorem 2, condition (i) can be relaxed. According to Scheffé's Lemma (see Lehmann, page 351, Lemma 4), a.e. convergence of densities implies convergence of the probability measures with respect to the supremum-metric (and hence also setwise and—if meaningful—weak and vague convergence). None of these weaker types of convergence is sufficient to guarantee a.e. convergence of the conditional expectations. For weak convergence this follows from an Example in Steck, page 238, for setwise convergence from Example 4.2 in Trumbo. Example 3 shows that not even convergence with respect to the supremum-metric (equivalently: convergence of densities in μ-mean or μ-measure) is sufficient.

Though these weaker types of convergence are not strong enough to imply a.e. convergence of the conditional expectations, they imply weaker types of convergence (see Theorem 3).

EXAMPLE 3. Let $X = [-1, 1]$, $\mathscr{F}$ the Borel sets in $[-1, 1]$ and $\lambda | \mathscr{F}$ the restriction of the Lebesgue measure. For any $n \in \mathbb{N}$ let $k_n \in \mathbb{N}$ and $l_n \in \{0, 1, \cdots, k_n - 1\}$ be such that $n = k_n(k_n - 1)/2 + l_n$. Let $A_n = [l_n k_n^{-1}, (l_n + 1)k_n^{-1}]$ and let $h_n(x) = (1 + (\text{sign } x) 1_{A_n}(|x|))/2$ and $h_0(x) \equiv \frac{1}{2}$. Let $P_n | \mathscr{F}$ be the probability measure having density h_n with respect to $\lambda | \mathscr{F}$, $n = 0, 1, 2, \cdots$. As $\{h_n : n = 0, 1, 2, \cdots\}$ is uniformly bounded by the function identically 1, Condition (ii) of Theorem 2 is fulfilled. Furthermore, $(P_n)_{n\in\mathbb{N}} \to P_0$ with respect to the supremum-metric. Let $\mathscr{F}_0$ be the sub-σ-field consisting of all sets in $\mathscr{F}$ which are symmetric about 0. It is easy to check that for each $n = 0, 1, 2, \cdots$

$$p_n(1_{[0,1]}, x) = h_n(|x|)$$

is a conditional expectation of $1_{[0,1]}$ relative to P_n, given $\mathscr{F}_0$. However $(p_n(1_{[0,1]}, \cdot))_{n\in\mathbb{N}}$ converges on a λ-null set only.

The following theorem generalizes and extends Satz (2.9) of Rhefus.

THEOREM 3. *Assume that*

(i) $(h_n)_{n\in\mathbb{N}}$ *converges to* h_0 *in* μ*-measure. Then*

(11) *For arbitrary versions of the conditional expectations*: $(p_n(f,\cdot))_{n\in\mathbb{N}}$ *converges to* $p_0(f,\cdot)$ *in* P_0*-measure for any* $\mathscr{F}$*-measurable, bounded function* f.

(12) *If* $\mathscr{F}$ *is countably generated and if a regular conditional probability relative to* μ, *given* $\mathscr{F}_0$, *exists, then* $(\sup_{A\in\mathfrak{F}}|p_n(A,x)-p_0(A,x)|)_{n\in\mathbb{N}}$ *converges to* 0 *in* P_0*-measure for all regular conditional probabilities* $p_n\,|\,\mathscr{F}\times X$ *relative to* P_n, *given* $\mathscr{F}_0$.

We remark that the condition of the densities depends neither on the particular dominating measure μ nor on the particular versions chosen. Hence w.l.g. $\mu(X)=1$.

We further remark that by Scheffé's Lemma convergence of the densities in μ-measure is equivalent to convergence of the measures with respect to the supremum-metric.

As f is bounded, $p_n(f,\cdot)$ is bounded P_n-a.e. For versions which are bounded *everywhere*, convergence in P_0-measure implies convergence in P_0-mean (see Halmos, page 110, Theorem D).

PROOF. By Scheffé's Lemma, $(h_n)_{n\in\mathbb{N}}\to h_0$ in μ-measure implies $(|h_n-h_0|)_{n\in\mathbb{N}}\to 0$ in μ-mean. Hence $(\mu(|h_n-h_0|,\cdot))_{n\in\mathbb{N}}\to 0$ in μ-mean and therefore $(\mu(|h_n-h_0|,\cdot))_{n\in\mathbb{N}}\to 0$ in μ-measure. As $|\mu(h_n,\cdot)-\mu(h_0,\cdot)|\leqq\mu(|h_n-h_0|,\cdot)$ μ-a.e., this implies that $(\mu(h_n,\cdot))_{n\in\mathbb{N}}$ converges to $\mu(h_0,\cdot)$ in μ-measure. Let $\mathbb{N}_0\subset\mathbb{N}$ be an arbitrary infinite subset and let $\mathbb{N}_1\subset\mathbb{N}_0$ be such that $(h_n)_{n\in\mathbb{N}_1}$ converges to h_0 μ-a.e. and $(\mu(h_n,\cdot))_{n\in\mathbb{N}_1}$ converges to $\mu(h_0,\cdot)$ μ-a.e. Since $\mu(h_n,\cdot)$ is a density of $P_n\,|\,\mathscr{F}_0$ with respect to $\mu\,|\,\mathscr{F}_0$, the assumptions of Theorem 1 are fulfilled with $\mathbb{N}$ replaced by $\mathbb{N}_1$. Theorem 1 (1) therefore implies that $(p_n(f,\cdot))_{n\in\mathbb{N}_1}$ converges to $p_0(f,\cdot)$ P_0-a.e. Hence any subsequence of $(p_n(f,\cdot))_{n\in\mathbb{N}}$ contains a subsequence converging to $p_0(f,\cdot)$ P_0-a.e. This, however, implies that $(p_n(f,\cdot))_{n\in\mathbb{N}}$ converges to $p_0(f,\cdot)$ in P_0-measure. This proves (11).

Let $p_n^*\,|\,\mathscr{F}\times X$ $n=0,1,2,\cdots$ be defined as in part (vi) of the proof of Theorem 1. As the parts (ii) and (iii) of the proof of Theorem 1 run through with $\mathbb{N}$ replaced by $\mathbb{N}_1$, one obtains

(13) $$\lim_{n\in\mathbb{N}_1}\sup_{A\in\mathscr{F}}|p_n^*(A,x)-p_0^*(A,x)|=0 \qquad \text{for} \quad P_0\text{-a.a. } x\in X.$$

As part (vii) of the proof of Theorem 1 also holds in the present case, we obtain together with (13)

$$\lim_{n\in\mathbb{N}_1}\sup_{A\in\mathscr{F}}|p_n(A,x)-p_0(A,x)|=0 \qquad \text{for} \quad P_0\text{-a.a. } x\in X$$

for all regular conditional probabilities $p_n\,|\,\mathscr{F}\times X$. As $\mathscr{F}$ is assumed to be countably generated, the functions $S_n(x)=\sup_{A\in\mathscr{F}}|p_n(A,x)-p_0(A,x)|$, $n\in\mathbb{N}_1$, are $\mathscr{F}_0$-measurable, since the suprema over $\mathscr{F}$ equal the suprema over $\mathscr{H}$, where $\mathscr{H}$ is any countable algebra generating $\mathscr{F}$.

Hence any subsequence of $(S_n)_{n\in\mathbb{N}}$ contains a subsequence converging to 0 P_0-a.e. This, however, implies that $(S_n)_{n\in\mathbb{N}}$ converges to 0 in P_0-measure. This proves (12).

We remark that Example 2 also shows that the boundedness assumed for f in assertion (11) of Theorem 3 is essential. Example (2.1) of Rhefus shows that setwise convergence of the probability measures is not sufficient to guarantee convergence of conditional expectations in any reasonable sense, in particular: setwise convergence of the probability measures is not sufficient to guarantee convergence of conditional expectations in P_0-measure.

For the sake of completeness we shall state another theorem which follows easily from Lemma 2. This theorem generalizes Theorem 5.6 of Trumbo (mainly by eliminating Trumbo's assumption that $\mathscr{F}_0$ is countably generated).

THEOREM 4. *Let* $f: X \to [0, 1]$ *be* $\mathscr{F}$*-measurable. Assume that the following condition is fulfilled for* $g = f$ *and* $g = 1-f$:

(14) $(P_n(g1_{A_0})/P_0(g1_{A_0}))_{n\in\mathbb{N}} \to 1$ *uniformly on the class of all* $A_0 \in \mathscr{F}_0$ *with* $P_0(g1_{A_0}) > 0$.

Then for arbitrary versions of the conditional expectations, $(p_n(f,\cdot))_{n\in\mathbb{N}}$ *converges to* $p_0(f,\cdot)$ P_0-a.e.

If $P_n|\mathscr{F} \ll P_0|\mathscr{F}$ *for all* $n\in\mathbb{N}$, *it suffices to require* (14) *for* $g = f$ *and* $g \equiv 1$.

PROOF. Let $\mu|\mathscr{F}$ be a probability measure dominating $P_n|\mathscr{F}$, $n = 0, 1, 2, \cdots$ and let h_n be a density of $P_n|\mathscr{F}$ with respect to $\mu|\mathscr{F}$. We shall show that the assumptions of Lemma 2 are fulfilled for f. This then immediately implies the assertion.

We remark that (14) for $g = f$ and $g = 1-f$ implies that (14) is also fulfilled for $g \equiv 1$. As $P_n(g1_{A_0}) = \mu(\mu(h_n g,\cdot)1_{A_0})$, the assumption of Lemma 3 with $g_n = \mu(h_n g,\cdot)$, $n = 0, 1, 2, \cdots$, is fulfilled for $g = f$, $g = 1-f$ and $g \equiv 1$. Hence there exists a μ-null set $N_g \in \mathscr{F}_0$ such that

(15) $x \in \overline{N}_g$ and $\mu(h_0 g, x) > 0$ implies $(\mu(h_n g, x))_{n\in\mathbb{N}} \to \mu(h_0 g, x)$.

Applied for $g = f$ this yields

$$(\mu(h_n f, x))_{n\in\mathbb{N}} \to \mu(h_0 f, x) \quad \text{for all} \quad x \in \overline{N}_f \quad \text{with} \quad \mu(h_0 f, x) > 0.$$

To obtain convergence also in the case $x \in \overline{N}_f$, $\mu(h_0 f, x) = 0$, we have to use that $\mu(h_0 f, x) = 0$ and $\mu(h_0, x) > 0$ together imply $\mu(h_0(1-f), x) > 0$ for μ-a.a. $x \in X$. Hence (by (15) applied for $g = 1-f$ and $g \equiv 1$) there exists a μ-null set $N_0 \supset N_f$ such that $x \in \overline{N}_0$, $\mu(h_0 f, x) = 0$ and $\mu(h_0, x) > 0$ implies

$$(\mu(h_n, x))_{n\in\mathbb{N}} \to \mu(h_0, x) \quad \text{and} \quad (\mu(h_n(1-f), x))_{n\in\mathbb{N}} \to \mu(h_0(1-f)).$$

Both relations together imply $(\mu(h_n f, x))_{n\in\mathbb{N}} \to \mu(h_0 f, x)$ for all $x \in \overline{N}_1$ with $\mu(h_0 f, x) = 0$ and $\mu(h_0, x) > 0$, where $N_1 \supset N_0$ and $P_0(N_1) = 0$. Hence we obtain for all $x \in \overline{N}_1$ with $\mu(h_0, x) > 0$,

$$(\mu(h_n, x))_{n\in\mathbb{N}} \to \mu(h_0, x) \quad \text{and} \quad (\mu(h_n f, x))_{n\in\mathbb{N}} \to \mu(h_0 f, x).$$

As $P_0\{x \in \overline{N}_1 : \mu(h_0, x) > 0\} = P_0\{x \in X : \mu(h_0, x) > 0\} = 1$, this implies that the assumptions of Lemma 2 are fulfilled for f.

If $P_n|\mathscr{F} \ll P_0|\mathscr{F}$, we have $\mu(h_n f,\cdot) = 0$ μ-a.e. on the set $\{x \in X : \mu(h_0 f, x) = 0\}$. Hence in this case (14) for $g = 1-f$ may be omitted.

The following corollary generalizes Satz (4.12) of Rhefus which refers to separable metric spaces.

COROLLARY 2. *Assume that condition* (14) *is fulfilled for all functions* $g = 1_B$, $B \in \mathscr{F}$. *Then for any* $\mathscr{F}$*-measurable, bounded function* f *and all versions of the conditional expectations*

$$(p_n(f, \cdot))_{n \in \mathbb{N}} \to p_0(f, \cdot) \; P_0\text{-a.e.}$$

PROOF. It suffices to prove the corollary for functions $f: X \to [0, 1]$. It is a matter of routine to show that (14) for all 1_B, $B \in \mathscr{F}$, implies (14) for all bounded non-negative $\mathscr{F}$-measurable functions g, hence in particular for $g = f$ and $g = 1 - f$. Thus the assertion follows from Theorem 4.

We remark that Example 2 also shows that the boundedness assumed for f in Corollary 2 is essential. (In this case (14) is fulfilled for every $g = 1_B$, $B \in \mathscr{F}$; but the conditional expectations do not converge a.e.)

2. A few lemmas.

LEMMA 1. *Let* $\mu | \mathscr{F}_0$ *be a* σ*-finite measure dominating* $P_n | \mathscr{F}_0$, $n = 0, 1, 2, \cdots$. *Let* h_{0n} *be a density of* $P_n | \mathscr{F}_0$ *with respect to* $\mu | \mathscr{F}_0$. *Assume that*

$$h_{00} \leqq \liminf_{n \in \mathbb{N}} h_{0n} \; P_0\text{-a.e.} \tag{16}$$

Then $P_0(\limsup_{n \in \mathbb{N}} (A_n \cup A_0)) = 0$ *for all sequences* $(A_n)_{n=0,1,2,\cdots}$ *with* $A_n \in \mathscr{F}_0$ *and* $P_n(A_n) = 0$ *for all* $n = 0, 1, 2, \cdots$.

PROOF. Let $H_n = \{x \in X: h_{0n}(x) = 0\}$. Then $x \in \limsup_{n \in \mathbb{N}} H_n$ implies $\liminf_{n \in \mathbb{N}} h_{0n}(x) = 0$.

As $h_{00}(x) \leqq \liminf_{n \in \mathbb{N}} h_{0n}(x)$ P_0-a.e., we have $h_{00}(x) = 0$ for P_0-a.a. $x \in \limsup_{n \in \mathbb{N}} H_n$, whence $P_0(\limsup_{n \in \mathbb{N}} H_n) = 0$. Let $(A_n)_{n=0,1,2,\cdots}$ be a sequence of sets in $\mathscr{F}_0$ with $P_n(A_n) = 0$ for $n = 0, 1, 2, \cdots$. Then $\mu(A_n - H_n) = 0$ and therefore $P_0(A_n - H_n) = 0$ for all $n = 0, 1, 2, \cdots$. As $P_0(\limsup_{n \in \mathbb{N}} H_n) = 0$, this implies $P_0(\limsup_{n \in \mathbb{N}} (A_n \cup A_0)) = 0$.

LEMMA 2. *Let* $\mu | \mathscr{F}$ *be a probability measure dominating* $P_n | \mathscr{F}$ *for* $n = 0, 1, 2, \cdots$. *If for some* $\mathscr{F}$*-measurable, bounded function* f

$$(\mu(h_n f, \cdot))_{n \in \mathbb{N}} \to \mu(h_0 f, \cdot) \; P_0\text{-a.e.}$$

and if, furthermore,

$$(\mu(h_n, \cdot))_{n \in \mathbb{N}} \to \mu(h_0, \cdot) \; P_0\text{-a.e.}$$

then $(p_n(f, \cdot))_{n \in \mathbb{N}} \to p_0(f, \cdot)$ P_0-a.e. *for arbitrary versions of the conditional expectations.*

PROOF. (i) At first we shall show that

$$\begin{aligned} p_n^*(f, x) &= \mu(h_n f, x)/\mu(h_n, x) && \text{if } \mu(h_n, x) > 0, \\ &= 0 && \text{otherwise}; \qquad n = 0, 1, 2, \cdots \end{aligned}$$

is a conditional expectation of f relative to P_n, given $\mathscr{F}_0$.

As $p_n^*(f, \cdot)$ is obviously $\mathscr{F}_0$-measurable, it remains to be shown that

$$P_n(p_n^*(f, \cdot) 1_{A_0}) = P_n(f 1_{A_0}) \qquad \text{for all } A_0 \in \mathscr{F}_0. \tag{17}$$

As $\mu(f,\cdot)1_{A_0} = \mu(f1_{A_0},\cdot)$ μ-a.e. and as $\mu(h_n,\cdot)$ is a density of $P_n|\mathscr{F}_0$ with respect to $\mu|\mathscr{F}_0$, (17) follows immediately. This proves that $p_n^*(f,\cdot)$ is a conditional expectation.

(ii) Now we shall show that the assertion holds for the particular versions $p_n^*(f,\cdot)$, $n = 0, 1, 2, \cdots$. Let N denote the P_0-null set comprising all those $x \in X$ for which $(\mu(h_n, x))_{n\in\mathbb{N}} \to \mu(h_0, x)$ or $(\mu(h_n f, x))_{n\in\mathbb{N}} \to \mu(h_0 f, x)$ fails to be true. $x \in \bar{N}$ and $\mu(h_0, x) > 0$ together imply $p_n^*(f,x) = \mu(h_n f, x)/\mu(h_n, x)$ for all sufficiently large n, whence $\lim_{n\in\mathbb{N}} p_n^*(f,x) = p_0^*(f,x)$. As $P_0(\bar{N}) = 1$, we have

$$P_0\{x\in\bar{N}: \mu(h_0,x) > 0\} = P_0\{x\in X: \mu(h_0,x)>0\} = 1.$$

(iii) Finally we shall show that the assertion holds for arbitrary versions $p_n(f,\cdot)$, $n = 0, 1, 2, \cdots$. This, however, follows immediately from Lemma 1 applied for $h_{0n} = \mu(h_n,\cdot)$ and $A_n = \{x\in X: p_n^*(f,x) \neq p_n(f,x)\}$, $n = 0, 1, 2, \cdots$. We have $A_n \in \mathscr{F}_0$ and $P_n(A_n) = 0$, $n = 0, 1, 2, \cdots$, whence $P_0(\limsup_{n\in\mathbb{N}}(A_n \cup A_0)) = 0$. Hence for P_0-a.a. $x \in X$ there exists $n(x)\in\mathbb{N}$ such that $p_0(f,x) = p_0^*(f,x)$ and $p_n(f,x) = p_n^*(f,x)$ for all $n \geqq n(x)$. This implies the assertion.

LEMMA 3. *Let g_n be $\mathscr{F}_0$-measurable for $n = 0, 1, 2, \cdots$. If*

$$(\mu(g_n 1_{A_0})/\mu(g_0 1_{A_0}))_{n\in\mathbb{N}} \to 1$$

uniformly on the class of all $A_0 \in \mathscr{F}_0$ with $\mu(g_0 1_{A_0}) > 0$, then $(g_n(x)/g_0(x))_{n\in\mathbb{N}} \to 1$ μ-a.e. uniformly on the sets of all $x\in X$ with $g_0(x) > 0$.

PROOF. Let $A_* = \{x \in X: g_0(x) > 0\}$, $\alpha_n = \sup\{\mu(g_n 1_{A_0}): A_0 \in \mathscr{F}_0, \mu(g_0 1_{A_0}) > 0\}$ and $\beta_n = \mu\text{-ess.sup}\{g_n(x)/g_0(x): x\in A_*\}$. For $r < \beta_n$ let $A_{n,r} = \{x \in A_*: (g_n(x)/g_0(x)) > r\}$. By definition of β_n, we have $\mu(A_{n,r}) > 0$. As $g_0(x) > 0$ for $x \in A_{n,r}$, this implies $\mu(g_0 1_{A_{n,r}}) > 0$. Hence $r < (\mu(g_n 1_{A_{n,r}})/\mu(g_0 1_{A_{n,r}})) \leqq \alpha_n$. As $r < \beta_n$ was arbitrary, this implies $\beta_n \leqq \alpha_n$. By assumption, $\lim_{n\in\mathbb{N}}\alpha_n = 1$. Hence $\limsup_{n\in\mathbb{N}}\beta_n \leqq 1$. For $s > \beta_n$ let $B_{n,s} = \{x\in A_*: (g_n(x)/g_0(x)) < s\}$. By definition of β_n we have $\mu(B_{n,s}) = 1$. As $g_0(x) > 0$ for $x \in B_{n,s}$, this implies $\mu(g_0 1_{B_{n,s}}) > 0$. Hence $(\mu(g_n 1_{B_{n,s}})/\mu(g_0 1_{B_{n,s}})) < s$ and therefore $\alpha_n \leqq s$. As $s > \beta_n$ was arbitrary, this implies $\alpha_n \leqq \beta_n$, hence $\liminf_{n\in\mathbb{N}}\beta_n \geqq \lim_{n\in\mathbb{N}}\alpha_n = 1$.

REFERENCES

[1] DOOB, J. L. (1953). *Stochastic Processes*. Wiley, New York.

[2] GIHMAN, I. I. (1953a). On some limit theorems for conditional distributions and on problems of mathematical statistics connected with them. *Ukrain. Mat. Z.* **5** 413–433 (in Russian). Eng. tr. *Selected Translations* **2** 5–26.

[3] GIHMAN, I. I. (1953b). Some limit theorems for conditional distributions. *DAN SSSR (N.S.)* **91** 1003–1006.

[4] HALMOS, P. R. (1964). *Measure Theory*. Van Nostrand, Princeton.

[5] LEHMANN, E. L. (1959). *Testing Statistical Hypotheses*. Wiley, New York.

[6] RHEFUS, H. H. (1969). Zur Konvergenz bedingter Wahrscheinlichkeiten. Diplomarbeit, Universität zu Köln.

[7] STECK, G. P. (1957). Limit theorems for conditional distributions. *Univ. of California Publ. Statist.* **2** 237–284.

[8] TRUMBO, B. E. (1965). Sufficient conditions for weak convergence of conditional probability distributions in a metric space. Ph.D. dissertation, Univ. of Chicago.