AN INEQUALITY FOR THE HAUSDORFF-METRIC
OF σ-FIELDS

By D. Landers and L. Rogge

University of Cologne and University-GH-Duisburg

It is shown that the Hausdorff-metric of σ-fields—which plays an important role for uniform martingale theorems—has a surprising "additivity" property. For example this property can be used to obtain a sharpened version of a uniform inequality for conditional expectations.

1. Introduction. Let $(\Omega, \mathscr{A}, P)$ be a probability space. To avoid proliferation of symbols we use the same symbol $\rho(\cdot, \cdot)$ for several distance functions. If $A \in \mathscr{A}$ and $\mathscr{B} \subset \mathscr{A}$ denote by $\rho(A, \mathscr{B}) = \inf_{B \in \mathscr{B}} P(A \triangle B) = \inf_{B \in \mathscr{B}} \|1_A - 1_B\|_1$ the $\| \ \|_1$-distance of A from $\mathscr{B}$. If $\mathscr{A}_0, \mathscr{B}_0 \subset \mathscr{A}$ we write

$$\rho(\mathscr{A}_0, \mathscr{B}_0) = \sup_{A \in \mathscr{A}_0} \rho(A, \mathscr{B}_0)$$

and

$$d(\mathscr{A}_0, \mathscr{B}_0) = \rho(\mathscr{A}_0, \mathscr{B}_0) + \rho(\mathscr{B}_0, \mathscr{A}_0).$$

Then d is a pseudometric on the set of all nonvoid subsystems of $\mathscr{A}$. If we endow $\mathscr{A}$ with the pseudometric $(A, B) \to P(A \triangle B)$, then all complete sub-σ-fields of $\mathscr{A}$ are closed subsets of $\mathscr{A}$ and d is equivalent to the usual Hausdorff-metric between closed subsets.

The pseudometric d was studied by Boylan (1971), Neveu (1972), Rogge (1974), Brunk (1975), and Mukerjee (1984). Boylan used this pseudometric to show that if a sequence of σ-fields $\mathscr{A}_n$ increases or decreases to a σ-field $\mathscr{A}_\infty$ and if $d(\mathscr{A}_n, \mathscr{A}_\infty) \to 0$ then

$$\delta_n = \sup_{f \in \Phi} \|P^{\mathscr{A}_n} f - P^{\mathscr{A}_\infty} f\|_1 \to 0,$$

where $P^{\mathscr{B}}$ denotes the conditional expectation operator and Φ is the system of all $\mathscr{A}$-measurable functions with values in $[0, 1]$.

Neveu (1972), Rogge (1974), and Brunk (1975) gave rates of convergence for δ_n. Neveu proved that

$$\delta_n = O\big(d(\mathscr{A}_n, \mathscr{A}_\infty)\big)$$

if $\mathscr{A}_n$ are σ-fields increasing or decreasing to the σ-field $\mathscr{A}_\infty$. Rogge proved for arbitrary σ-fields (not necessarily ordered by inclusion)

$$\delta_n = O\big(d(\mathscr{A}_n, \mathscr{A}_\infty)^{1/2}\big).$$

Brunk proved for arbitrary σ-lattices

$$\delta_n = O\!\left(d(\mathcal{A}_n, \mathcal{A}_\infty)^{1/4}\right).$$

The cited results of Brunk and Rogge follow from their $\|\ \|_2$-inequalities. In this paper it is shown that

$$\delta_n = O\!\left(d(\mathcal{A}_n, \mathcal{A}_\infty)\right)$$

for arbitrary σ-fields (see Corollary 5). It is an open question whether this holds for σ-lattices, too.

Let us remark that $O(d(\mathcal{A}_n, \mathcal{A}_\infty))$ is a sharp convergence rate for δ_n since $\delta_n \geq \frac{1}{2}d(\mathcal{A}_n, \mathcal{A}_\infty)$ always (see Rogge, 1974, page 489). Corollary 5 is a consequence of our inequality $\sum_{i \in N}\rho(A_i, \mathcal{B}_0) \leq 4\rho(\mathcal{A}_0, \mathcal{B}_0)$, if $A_i \in \mathcal{A}_0$ are disjoint (see Theorem 1). The distances $\rho(A, \mathcal{B}_n)$ play an important role for convergence orders in the conditional central limit theorem of Rényi (see Landers and Rogge, 1984a). The inequality above is one of the basic tools to obtain rates of convergence in the central limit theorem for sums of a random number τ_n of independent terms where τ_n/n converges to a nonconstant limit function τ (see Landers and Rogge, 1984b). Hitherto rates of convergence were known only for constant τ (see Landers and Rogge, 1976, 1977, and the literature cited there).

2. The results. The following inequality is the main tool of this paper and may be of independent interest.

THEOREM 1.   *Let $\mathcal{A}_0$, $\mathcal{B}_0 \subset \mathcal{A}$ be σ-fields and let $A_i \in \mathcal{A}_0$, $i \in I$, be disjoint sets. Then*

$$\sum_{i \in I} \rho(A_i, \mathcal{B}_0) \leq 4\rho(\mathcal{A}_0, \mathcal{B}_0).$$

PROOF.   Obviously it suffices to prove the assertion for the case that $I = \{1, \ldots, n\}$, $n \geq 3$, and $\sum_{i=1}^n A_i = \Omega$. According to Kudo (1974, Lemma 2.1) there holds for each $A \in \mathcal{A}$

(1) $\rho(A, \mathcal{B}_0) = \frac{1}{2} - E\!\left(|\frac{1}{2} - P^{\mathcal{B}_0}A|\right) = E\!\left(P^{\mathcal{B}_0}A \wedge (1 - P^{\mathcal{B}_0}A)\right).$

Let $\varphi_i = P^{\mathcal{B}_0}A_i$, $i \in I$, and put for each $N \subset I$

$$\varphi_N = \sum_{i \in N} \varphi_i, \qquad A_N = \sum_{i \in N} A_i.$$

By (1) we have for each $N \subset I$

(2) $\rho(A_N, \mathcal{B}_0) = E\!\left(\varphi_N \wedge (1 - \varphi_N)\right).$

Choose $M \subset I$ such that

(3) $\rho(A_N, \mathcal{B}_0) \leq \rho(A_M, \mathcal{B}_0)$ for all $N \subset I$.

Since $\rho(A_M, \mathcal{B}_0) \le \rho(\mathcal{A}_0, \mathcal{B}_0)$ it suffices to show according to (2) that

$$(4) \qquad E\left(\sum_{i=1}^{n} \varphi_i \wedge (1 - \varphi_i) \right) \le 4E(\varphi_M \wedge (1 - \varphi_M)).$$

Since $|I| = n$ we have by (2) and (3)

$$E\left(\sum_{N \subset I} \varphi_N \wedge (1 - \varphi_N) \right) \le 2^n E(\varphi_M \wedge (1 - \varphi_M)).$$

Hence (4) is shown if we prove

$$(5) \qquad (2^n/4) \sum_{i=1}^{n} \varphi_i \wedge (1 - \varphi_i) \le \sum_{N \subset I} \varphi_N \wedge (1 - \varphi_N) \quad \text{on} \left\{ \sum_{i=1}^{n} \varphi_i = 1 \right\}.$$

Consider the following two cases,

(A) $\varphi_{i_0}(\omega) > \frac{1}{2}$ for some $i_0 \in \{1, \dots, n\}$ and
(B) $\varphi_i(\omega) \le \frac{1}{2}$ for all $i \in \{1, \dots, n\}$, $\omega \in \{\sum_{i=1}^{n}\varphi_i = 1\}$,

and write φ_i instead of $\varphi_i(\omega)$.

CASE (A).   Since $\sum_{i=1}^{n}\varphi_i = 1$ we have

$$(6) \qquad \sum_{i=1}^{n} \varphi_i \wedge (1 - \varphi_i) = 1 - \varphi_{i_0} + \sum_{i \ne i_0} \varphi_i = 2(1 - \varphi_{i_0}).$$

Furthermore we have

$$\begin{aligned}
\sum_{N \subset I} \varphi_N \wedge (1 - \varphi_N) &= \sum_{i_0 \in N \subset I} \varphi_N \wedge (1 - \varphi_N) + \sum_{i_0 \notin N \subset I} \varphi_N \wedge (1 - \varphi_N) \\
&= \sum_{i_0 \in N \subset I} (1 - \varphi_N) + \sum_{i_0 \notin N \subset I} \varphi_N \\
&= \sum_{i_0 \notin N \subset I} (1 - (\varphi_N + \varphi_{i_0})) + \sum_{i_0 \notin N \subset I} \varphi_N \\
&= \sum_{i_0 \notin N \subset I} (1 - \varphi_{i_0}) = \tfrac{1}{2} 2^n (1 - \varphi_{i_0}).
\end{aligned}$$

Together with (6) this implies (5).

CASE (B).   In this case $\sum_{i=1}^{n}\varphi_i \wedge (1 - \varphi_i) = \sum_{i=1}^{n}\varphi_i = 1$ and hence we have to show that

$$(7) \qquad \sum_{N \subset I} \varphi_N \wedge (1 - \varphi_N) \ge 2^n/4.$$

We prove this inductively for $n \ge 3$. Let $n = 3$, i.e., $I = \{1, 2, 3\}$. Since $\varphi_1 + \varphi_2 + \varphi_3 = 1$ and $\varphi_i \le \frac{1}{2}$, $i \in I$, we have $\varphi_i + \varphi_j \ge \frac{1}{2}$ for $i \ne j$ and (7) holds with equality. Now assume that (7) holds for $n \ge 3$ (i.e., (7) holds for all real numbers $\varphi_i \in [0, \frac{1}{2}]$ with $\sum_{i=1}^{n}\varphi_i = 1$).

Let $\varphi_1, \dots, \varphi_{n+1} \in [0, \frac{1}{2}]$ with $\sum_{i=1}^{n+1}\varphi_i = 1$. Since $n \ge 3$, w.l.g. we may assume $\varphi_n + \varphi_{n+1} \le \frac{1}{2}$. Put $\psi_i = \varphi_i$ for $i = 1, \dots, n-1$ and $\psi_n = \varphi_n + \varphi_{n+1}$. By induc-

tive assumption we have

$$\sum_{N \subset \{1,\dots,n\}} \psi_N \wedge (1 - \psi_N) \geq 2^n/4.$$

Hence it suffices to prove

$$(8) \qquad \sum_{N \subset \{1,\dots,n+1\}} \varphi_N \wedge (1 - \varphi_N) \geq 2 \sum_{N \subset \{1,\dots,n\}} \psi_N \wedge (1 - \psi_N).$$

Obviously (8) is shown if we prove

$$(9) \qquad \begin{aligned} \sum_{N \subset \{1,\dots,n-1\}} &\left(\varphi_{N+\{n\}} \wedge \left(1 - \varphi_{N+\{n\}}\right) + \varphi_{N+\{n+1\}} \wedge \left(1 - \varphi_{N+\{n+1\}}\right) \right) \\ &\geq \sum_{N \subset \{1,\dots,n-1\}} \left(\psi_N \wedge (1 - \psi_N) + \psi_{N+\{n\}} \wedge \left(1 - \psi_{N+\{n\}}\right) \right). \end{aligned}$$

A direct computation shows that for all $a, b, c \in [0,1]$ with $a + b + c \leq 1$, $b + c \leq \frac{1}{2}$ there holds

$$(10) \qquad \begin{aligned} (a + b) \wedge &\left(1 - (a + b)\right) + (a + c) \wedge \left(1 - (a + c)\right) \\ &\geq a \wedge (1 - a) + (a + b + c) \wedge \left(1 - (a + b + c)\right). \end{aligned}$$

Now let $N \subset \{1,\dots,n-1\}$ be fixed and put $a = \psi_N$, $b = \varphi_n$, and $c = \varphi_{n+1}$. Then $\varphi_{N+\{n\}} = a + b$, $\varphi_{N+\{n+1\}} = a + c$, $\psi_{N+\{n\}} = a + b + c$, and an application of (10) yields (9).

The following example shows that the constant 4, appearing in the inequality of Theorem 1, cannot be replaced by a constant less than 3.5. By a rather technical and tedious modification of this example, it can be seen that also the constant 3.5 does not work. We believe that the constant 4 is optimal.

EXAMPLE 2. We construct a probability space $(\Omega, \mathscr{A}, P)$, a σ-field $\mathscr{B}_0 \subset \mathscr{A}$ and disjoint $A_i \in \mathscr{A}$, $i \in I = \{1,2,\dots,7\}$, $\sum_{i \in I} A_i = \Omega$, such that with the σ-field $\mathscr{A}_0 := \{A_N = \sum_{i \in N} A_i : N \subset I\}$

$$(11) \qquad \sum_{i \in I} \rho(A_i, \mathscr{B}_0) = 1,$$

$$(12) \qquad \rho(\mathscr{A}_0, \mathscr{B}_0) = \sup_{N \subset I} \rho(A_N, \mathscr{B}_0) = \tfrac{2}{7}.$$

Let $X = \{1,2,3\}$, $Y = \{1,2,\dots,7\}$, and put $\Omega = X \times Y$. Let $\mathscr{A}$ be the power set of Ω, $\mathscr{B}_0 = \{X \times B : B \subset Y\}$, and $P|\mathscr{A}$ be the p-measure, defined by $P\{\omega\} = \frac{1}{21}$, for all $\omega \in \Omega$. To define the set A_i put at first $B_1 = \{1,2,3\}$, $B_2 = \{1,4,5\}$, $B_3 = \{1,6,7\}$, $B_4 = \{2,4,6\}$, $B_5 = \{2,5,7\}$, $B_6 = \{3,4,7\}$, and $B_7 = \{3,5,6\}$.

Then $B_i \subset Y$, $i \in I = \{1,\dots,7\}$, and we have

$$(13) \qquad \sum_{i \in I} 1_{B_i}(y) = 3 \quad \text{for all } y \in Y;$$

$$(14) \qquad \sup_{N \in I} \# \left\{ y \in Y : \sum_{i \in N} 1_{B_i}(y) \in \{1,2\} \right\} = 6.$$

By (13) for each $y \in Y$ there exist unique $1 \leq y(1) < y(2) < y(3) \leq 7$ such that $y \in B_{y(j)}$, $j = 1,2,3$.

Now define
$$A_i = \{(x, y) \in \Omega;\ y \in B_i,\ y(x) = i\}, \qquad i \in I.$$
These $A_i \in \mathscr{A}$, $i \in I$, are disjoint and $\sum_{i \in I} A_i = \Omega$.

We directly obtain

(15) $$\varphi_i = P^{\mathscr{B}_0} A_i = (1/3) 1_{X \times B_i}, \qquad i \in I.$$

Now (15) implies (use (2))
$$\sum_{i \in I} \rho(A_i, \mathscr{B}_0) = \sum_{i \in I} E(\varphi_i \wedge (1 - \varphi_i)) = \sum_{i \in I} E(\varphi_i) = 1,$$

i.e., (11) is fulfilled. For each $N \subset I$ we have by (15) that
$$\varphi_N \wedge (1 - \varphi_N) \in \left\{0, \tfrac{1}{3}\right\}$$

and
$$\varphi_N \wedge (1 - \varphi_N)(x, y) = \tfrac{1}{3} \quad \text{iff} \quad \sum_{i \in N} 1_{B_i}(y) \in \{1, 2\}.$$

Hence (14) implies (use (2))
$$\rho(\mathscr{A}_0, \mathscr{B}_0) = \sup_{N \subset I} E(\varphi_N \wedge (1 - \varphi_N))$$
$$= \sup_{N \subset I} \left(\tfrac{1}{21}\right)\left(\tfrac{1}{3}\right)(\#X)\#\left\{y \in Y: \sum_{i \in N} 1_{B_i}(y) \in \{1, 2\}\right\} = \tfrac{2}{7},$$

i.e., (12) is fulfilled.

For an integrable function f let
$$\rho(f, \mathscr{B}_0) = \inf\{\|f - g\|_1 : g \text{ is } \mathscr{B}_0\text{-measurable}\}.$$

Then Theorem 1 yields also a result for measurable functions instead of indicator functions.

COROLLARY 3.   *Let $\mathscr{A}_0$, $\mathscr{B}_0 \subset \mathscr{A}$ be σ-fields and let $f_i \geq 0$ be $\mathscr{A}_0$-measurable functions with $\sum_{i=1}^n f_i \leq 1$. Then*
$$\sum_{i=1}^n \rho(f_i, \mathscr{B}_0) \leq 4\rho(\mathscr{A}_0, \mathscr{B}_0).$$

PROOF.   W.l.g. we may assume that $f_1, \ldots, f_n$ are $\mathscr{A}_0$-measurable step functions. Hence there exists a common representation
$$f_i = \sum_{\nu=1}^k \alpha_{i\nu} 1_{A_\nu}, \qquad i = 1, \ldots, n,$$

where $A_1, \ldots, A_k \in \mathscr{A}_0$ are disjoint. As $\alpha_{i\nu} \geq 0$ and $\sum_{i=1}^n \alpha_{i\nu} \leq 1$, we obtain
$$\sum_{i=1}^n \rho(f_i, \mathscr{B}_0) \leq \sum_{i=1}^n \sum_{\nu=1}^k \alpha_{i\nu} \rho(A_\nu, \mathscr{B}_0)$$
$$\leq \sum_{\nu=1}^k \rho(A_\nu, \mathscr{B}_0) \leq 4\rho(\mathscr{A}_0, \mathscr{B}_0),$$

where the last inequality follows from Theorem 1.

COROLLARY 4. *Let $\mathscr{A}_0, \mathscr{B}_0 \subset \mathscr{A}$ be σ-fields and denote by $\mathscr{A}_0 \vee \mathscr{B}_0$ the smallest σ-field containing $\mathscr{A}_0$ and $\mathscr{B}_0$. Then*

$$\rho(\mathscr{A}_0 \vee \mathscr{B}_0, \mathscr{B}_0) \le 4\rho(\mathscr{A}_0, \mathscr{B}_0).$$

PROOF. It is easy to see that the system

$$\mathscr{C} = \left\{ \bigcup_{i=1}^{n} (A_i \cap B_i) : A_1, \ldots, A_n \in \mathscr{A}_0 \text{ disjoint}, B_1, \ldots, B_n \in \mathscr{B}_0, n \in \mathbb{N} \right\}$$

is a field, generating $\mathscr{A}_0 \vee \mathscr{B}_0$. Hence it suffices to show that

$$\rho(C, \mathscr{B}_0) \le 4\rho(\mathscr{A}_0, \mathscr{B}_0) \quad \text{for all } C \in \mathscr{C}.$$

Let $C = \bigcup_{i=1}^{n}(A_i \cap B_i) \in \mathscr{C}$. As $A_1, \ldots, A_n \in \mathscr{A}_0$ are disjoint and $B_1, \ldots, B_n \in \mathscr{B}_0$ we obtain

$$\rho(C, \mathscr{B}_0) \le \sum_{i=1}^{n} \rho(A_i \cap B_i, \mathscr{B}_0) \le \sum_{i=1}^{n} \rho(A_i, \mathscr{B}_0) \le 4\rho(\mathscr{A}_0, \mathscr{B}_0),$$

where the last inequality follows from Theorem 1.

COROLLARY 5. *Let $\mathscr{A}_0, \mathscr{B}_0 \subset \mathscr{A}$ be σ-fields. Then*

$$\sup_{f \in \Phi} \| P^{\mathscr{A}_0} f - P^{\mathscr{B}_0} f \|_1 \le 8d(\mathscr{A}_0, \mathscr{B}_0).$$

PROOF. Using Theorem 2 of Rogge (1974) and Corollary 4 we obtain for all $f \in \Phi$

$$\| P^{\mathscr{A}_0} f - P^{\mathscr{B}_0} f \|_1 \le \| P^{\mathscr{A}_0} f - P^{\mathscr{A}_0 \vee \mathscr{B}_0} f \|_1 + \| P^{\mathscr{A}_0 \vee \mathscr{B}_0} f - P^{\mathscr{B}_0} f \|_1$$

$$\le 2\rho(\mathscr{B}_0 \vee \mathscr{A}_0, \mathscr{A}_0) + 2\rho(\mathscr{A}_0 \vee \mathscr{B}_0, \mathscr{B}_0)$$

$$\le 8\rho(\mathscr{B}_0, \mathscr{A}_0) + 8\rho(\mathscr{A}_0, \mathscr{B}_0) = 8d(\mathscr{A}_0, \mathscr{B}_0).$$

If $\mathscr{A}_0 \subset \mathscr{B}_0$, the inequality of Corollary 5 is due to Neveu (1972). Up to now, without the restriction $\mathscr{A}_0 \subset \mathscr{B}_0$, there were known inequalities in terms of $d(\mathscr{A}_0, \mathscr{B}_0)$ for $\sup_{f \in \Phi} \| P^{\mathscr{A}_0} f - P^{\mathscr{B}_0} f \|_2$ only (see Brunk, 1975, and Rogge, 1974). The methods, however, use Hilbert space properties of L_2 and cannot be applied to obtain sharp bounds for $\sup_{f \in \Phi} \| P^{\mathscr{A}_0} f - P^{\mathscr{B}_0} f \|_1$. If we use $\| f \|_1 \le \| f \|_2$ then Rogge's result yields

$$\sup_{f \in \Phi} \| P^{\mathscr{A}_0} f - P^{\mathscr{B}_0} f \|_1 \le c\left(d(\mathscr{A}_0, \mathscr{B}_0)^{1/2} \right)$$

for arbitrary σ-fields $\mathscr{A}_0, \mathscr{B}_0$ and Brunk's result yields

$$\sup_{f \in \Phi} \| P^{\mathscr{A}_0} f - P^{\mathscr{B}_0} f \|_1 \le c\left(d(\mathscr{A}_0, \mathscr{B}_0)^{1/4} \right)$$

for arbitrary σ-lattices $\mathscr{A}_0, \mathscr{B}_0$ and some constant $c > 0$.

The following corollary is a direct consequence of Corollary 5. It sharpens Theorem 2.7 of Mukerjee (1984).

COROLLARY 6.   *Let $\mathscr{A}_n$, $n \in \mathbb{N} \cup \{\infty\}$ be σ-fields with $\sum_{n \in \mathbb{N}} d(\mathscr{A}_n, \mathscr{A}_\infty) < \infty$. Then $\sup_{f \in \Phi} E[\sup_{n \geq m} |P^{\mathscr{A}_n} f - P^{\mathscr{A}_\infty} f|] \to_{m \to \infty} 0$.*

PROOF.   For all $f \in \Phi$ we have by Corollary 5

$$E\left[\sup_{n \geq m} |P^{\mathscr{A}_n} f - P^{\mathscr{A}_\infty} f| \right] \leq \sum_{n \geq m} E\left[|P^{\mathscr{A}_n} f - P^{\mathscr{A}_\infty} f| \right] \leq 8 \sum_{n \geq m} d(\mathscr{A}_n, \mathscr{A}_\infty) \to_{m \to \infty} 0.$$

REFERENCES

BOYLAN, E. S. (1971). Equi-convergence of martingales. *Ann. Math. Statist.* **42** 552–559.

BRUNK, H. D. (1975). Uniform inequalities for conditional p-means given σ-lattices. *Ann. Probab.* **3** 1025–1030.

KUDO, H. (1974). A note on the strong convergence of σ-algebras. *Ann. Probab.* **2** 76–83.

LANDERS, D. and ROGGE, L. (1976). The exact approximation order in the central-limit-theorem for random summation. *Z. Wahrsch. verw. Gebiete* **36** 269–283.

LANDERS, D. and ROGGE, L. (1977). A counterexample in the approximation theory of random summation. *Ann. Probab.* **5** 1018–1023.

LANDERS, D. and ROGGE, L. (1984a). Exact approximation orders in the conditional central-limit-theorem. *Z. Wahrsch. verw. Gebiete* **66** 227–224.

LANDERS, D. and ROGGE, L. (1984b). Sharp rates of convergence in the random central limit theorem with non constant limit function. Submitted for publication.

MUKERJEE, H. G. (1984). Almost sure equiconvergence of conditional expectations. *Ann. Probab.* **12** 733–741.

NEVEU, J. (1972). Note on the tightness of the metric on the set of complete sub-σ-algebras of a probability space. *Ann. Math. Statist.* **43** 1369–1371.

ROGGE, L. (1974). Uniform inequalities for conditional expectations. *Ann. Probab.* **2** 486–489.

MATHEMATICAL INSTITUTE   FACHBEREICH 11
UNIVERSITY OF COLOGNE   UNIVERSITY-GH-DUISBURG
WEYERTAL 86-90   LOTHARSTRASSE 65
D-5000 COLOGNE   D-4100 DUISBURG
WEST GERMANY   WEST GERMANY