Probab. Theory Relat. Fields 106, 371–377 (1996)

Probability Theory and Related Fields

A note on nearby variables with nearby conditional laws

M. Römersperger

Institut für Mathematische Stochastik, Universität Freiburg, D-79104 Freiburg, Germany

Received: 21 June 1994 / In revised form: 3 April 1996

Summary. We apply a version of the Strassen–Dudley theorem for Markov kernels to obtain a strong approximation theorem for sequences of random variables with values in Polish spaces. The conditions are expressed in terms of the Prohorov distance of regular conditional distributions. By avoiding discrete approximation of the random variables we can improve the upper bounds for the approximation.

Mathematics Subject Classification (1991): 60F15, 60B05, 60F17

1. Introduction

Strong approximation results have been proved under various assumptions. These can be traced back to the work of Strassen (1964), where an almost sure invariance principle was established for sums of independent identically distributed random variables using the Skorohod embedding theorem. By the technique developed in Berkes and Philipp (1979) it was possible to derive invariance principles not only for real valued but also for random variables with values in a separable Banach space. Several papers deal with this situation (see Morrow and Philipp (1982), Monrad and Philipp (1991) or Philipp (1986)). This method also can be used to derive strong approximation theorems for time continuous stochastic processes (see Eberlein (1989, 1992) and Besdziek (1991)).

All these results are based on the following situation: Let X, Y be random variables on $(\Omega, \mathfrak{F}, P)$ with values in a Polish space S. Let μ denote a law on $S \times S$ with first marginal $\mathfrak{L}(Y)$ such that the second marginal is close to $\mathfrak{L}(X)$ (in some sense). Then, we are interested in the existence of a random variable Z on $(\Omega, \mathfrak{F}, P)$ such that (Y,Z) is distributed according to μ and Z is close to X with probability close to 1.

Present address: Stadtstraße 46, D-79104 Freiburg, Germany.

This procedure can be splitted into two steps: First, we look for an appropriate law on the product space and second, we have to find random variables defined on $(\Omega, \mathfrak{F}, P)$, which are distributed according to this law. The latter point is answered almost completely by Dudley and Philipp (1983), Lemma 2.11. For the first step, it is essential to find a proper formulation for the condition on the closeness of the conditional laws.

The purpose of this paper is to state the conditions for nearby variables in terms of a modification of the Prohorov distance of regular conditional distributions. The advantage of this is twofold: Avoiding the necessity of discrete approximation the proof becomes much more elegant and above this, we get a stronger quantitative estimate (see Remark 2.3). The simple and useful tool for this is the formulation of the Strassen-Dudley theorem for Markov kernels in Lemma 3.1.

We use this estimate to establish a strong approximation theorem for sequences of random variables where an appropriate adapted sequence of random variables will be defined on a given filtered space. This result can be applied to the approximation of stochastic processes (see Römersperger (1993) and Eberlein and Römersperger (1996)).

2. Statement of results

Let (S, σ_S) be a Polish space provided by the Borel σ-field $\mathfrak{B}(S)$ and denote by $\mathfrak{M}(S)$ the set of Borel probability measures on S. For $\mu, \nu \in \mathfrak{M}(S)$ and $\alpha > 0$ we set

$$\pi(\alpha, \mu, \nu) := \inf\{\beta > 0 | \mu(A) \leq \nu(A^{\alpha]}) + \beta \text{ for all closed A} \subset S\} ,$$

where $A^{\alpha]}$ denotes the closed α-neighbourhood of A. We denote by $\mathscr{M}(\mu, \nu)$ the set of measures $\lambda \in \mathfrak{M}(S \times S)$ with marginals μ, ν. We use the corresponding notation for the Polish spaces (R, σ_R) and (T, σ_T) and denote by $\mathfrak{L}(Y)$ or $Y(Q)$ the distribution of a random variable Y (under Q). $\sigma(Y)$ stands for the sub-σ-field generated by the random variable Y.

Our main result is the following:

Theorem 1. *Let $(\Omega, \mathfrak{F}, P)$, $(\mathscr{E}, \mathfrak{G}, Q)$ be probability spaces and let $\mathfrak{F}_1$, $\mathfrak{F}_2$ be sub-σ-fields of $\mathfrak{F}$ such that $\mathfrak{F}_1 \subset \mathfrak{F}_2$. We consider the following random variables, each having values in a Polish space:*

$$(V, Y, Z) : (\mathscr{E}, \mathfrak{G}, Q) \to (R \times S \times T, \mathfrak{B}(R \times S \times T)) ,$$

$$Y^* : (\Omega, \mathfrak{F}_1, P) \to (S, \mathfrak{B}(S)) ,$$

$$X : (\Omega, \mathfrak{F}_2, P) \to (T, \mathfrak{B}(T)) .$$

Assume that $\mathfrak{L}(Y) = \mathfrak{L}(Y^*)$ holds. For $\alpha > 0$ let $\beta : S \to [0, \infty[$ denote a measurable function such that there are regular versions of the conditional distributions for which

$$\pi(\alpha, P[X \in \cdot | Y^* = y], Q[Z \in \cdot | Y = y]) \leq \beta(y) \tag{2.1}$$

holds for all $y \in S$. *Assume that there exist a random variable* $U : (\Omega, \mathfrak{F}_2, P) \to \mathbb{R}$ *and a regular version of the conditional distribution* $P[U \in du | (X, Y^*) = (x, y)]$ *having no atoms for all* $(x, y) \in T \times S$. *Then there exist random variables*

$$Z^* : (\Omega, \mathfrak{F}_2, P) \to (T, \mathfrak{B}(T)) ,$$

$$V^* : (\Omega, \mathfrak{F}_2, P) \to (R, \mathfrak{B}(R)) ,$$

with the following properties:

$$\mathfrak{L}(Y, Z, V) = \mathfrak{L}(Y^*, Z^*, V^*) ,$$

$$P[\sigma_T(X, Z^*) > \alpha] \leq E_Q[\beta(Y)] .$$

Remark 2.1. At a first reading the random variables V and V^* can be neglected. They are of use in cases, where we want to redefine besides Z some other random variable V in keeping with the common distribution.

Remark 2.2. For the applicability of Theorem 1 it is crucial to express Ineq. (2.1) in terms of standard conditions. In the special case $T = \mathbb{R}^d$ this can be done using conditional characteristic functions. For $d = 1$ let N_0, ϱ be fixed with $2N_0 \geq \varrho$. $N(0, 1)$ denotes the standard normal distribution. Then a function $\beta(\cdot)$ satisfying Ineq. (2.1) can be chosen such that

$$\begin{aligned} & E_Q[\beta(Y)] \\ & \leq \quad 2\pi^{-1} N_0 \alpha \varrho^{-1} \int\limits_{\{|u| \leq 2\alpha^{-1}\varrho^2\}} E_Q\left[\left| E_Q[\exp(iuZ) - E_P[\exp(iuX)|Y^* = \cdot](Y)|\sigma(Y)]\right|\right] du \\ & \quad + P[|X| > N_0 \alpha \varrho^{-1}] + (7N_0 + 3) N(0, 1)[| \cdot | > \varrho] . \end{aligned}$$

By reason of symmetry the term $P[|X| > N_0\alpha\varrho^{-1}]$ on the right-hand side can be replaced by $Q[|Z| > N_0\alpha\varrho^{-1}]$. If X is independent of $\mathfrak{F}_1$ then we have a. s.

$$E_P[\exp(iuX)|Y^* = \cdot] = E_P[\exp(iuX)] .$$

Remark 2.3. To illustrate the use, let us mention that Theorem 1 in Monrad and Philipp (1991) can be deduced as a corollary of our main theorem with a better estimation, i.e. (for $d = 1$) the constant α can be chosen as

$$\alpha := 12T^{-1} \log(T) + 2\pi^{-1} \lambda T^2 + \delta$$

instead of (2.1.6) in Monrad and Philipp (1991). The same improvement holds for all other comparable estimates (e.g. in Berkes and Philipp (1979); Monrad and Philipp (1992)).

Theorem 2. *Let $(\mathscr{E}, \mathscr{G}, (\mathscr{G}_k)_{k\in\mathbb{N}}, Q)$ be a filtered space, i.e. a probability space provided by an increasing sequence of sub-σ-fields. Let $(V_k)_{k\in\mathbb{N}}$ and $(Z_k)_{k\in\mathbb{N}}$ be sequences of random variables defined on $(\mathscr{E}, \mathscr{G}, Q)$ and adapted to $(\mathscr{G}_k)_{k\in\mathbb{N}}$ such that V_k (resp. Z_k) takes values in a Polish space (R_k, σ_{R_k}) (resp. (T_k, σ_{T_k})) for each $k \in \mathbb{N}$. Let $(X_k)_{k\in\mathbb{N}}$ be an adapted sequence of random variables on the filtered space $(\Omega, \mathfrak{F}, (\mathfrak{F}_k)_{k\in\mathbb{N}}, P)$ such that X_k is independent of $\mathfrak{F}_{k-1}$ and takes values in (T_k, σ_{T_k}). Denote by $\mathfrak{F}_0$ and $\mathscr{G}_0$ the trivial σ-field. Assume that there exists an adapted sequence $(U_k)_{k\in\mathbb{N}}$ of real-valued random variables on $(\Omega, \mathfrak{F}, (\mathfrak{F}_k)_{k\in\mathbb{N}}, P)$ such that a suitable regular version $P[U_k \in \cdot | X_k = x]$ of the conditional distribution has no atoms for all $x \in T_k$ and that (U_k, X_k) is independent of $\mathfrak{F}_{k-1}$.*

Let $\beta_k : \prod_{j=1}^{k-1} R_j \times \prod_{j=1}^{k-1} T_j \to [0,1]$ be a measurable function such that for a fixed sequence $(\alpha_k)_{k\in\mathbb{N}}$ of non-negative numbers and for all $y \in \prod_{j=1}^{k-1} R_j \times \prod_{j=1}^{k-1} T_j$, $k \in \mathbb{N}$

$$\pi(\alpha_k, P[X_k \in \cdot], Q[Z_k \in \cdot | ((V_j)_{j\leq k-1}, (Z_j)_{j\leq k-1}) = y]) \leq \beta_k(y) \ . \tag{2.2}$$

Then there exist adapted sequences $(V_k^*)_{k\in\mathbb{N}}$, $(Z_k^*)_{k\in\mathbb{N}}$ of random variables on $(\Omega, \mathfrak{F}, (\mathfrak{F}_k)_{k\in\mathbb{N}}, P)$ with $\mathfrak{L}((V_k, Z_k)_{k\in\mathbb{N}}) = \mathfrak{L}((V_k^*, Z_k^*)_{k\in\mathbb{N}})$ such that for all $k \in \mathbb{N}$

$$P[\sigma_{T_k}(X_k, Z_k^*) > \alpha_k] \leq E_Q[\beta((V_j)_{j\leq k-1}, (Z_j)_{j\leq k-1})] \ .$$

Remark 2.4. Setting $Y_k := ((V_j)_{j\leq k-1}, (Z_j)_{j\leq k-1})$ and $Y_k^* := ((V_j^*)_{j\leq k-1}, (Z_j^*)_{j\leq k-1})$ for each $k \geq 2$ the proof follows inductively by application of Theorem 1.

Remark 2.5. In the special case $T_k = \mathbb{R}$ a sequence $(\beta_k(\cdot))_{k\in\mathbb{N}}$ of measurable functions satisfying (2.2) can be chosen such that for all non-negative numbers N_k, ϱ_k with $2N_k \geq \varrho_k$

$$\begin{aligned} & E_Q[\beta_k((V_j)_{j\leq k-1}, (Z_j)_{j\leq k-1})] \\ \leq \quad & 2\pi^{-1} N_k \alpha_k \varrho_k{}^{-1} \int\limits_{\{|u|\leq 2\alpha_k{}^{-1}\varrho_k{}^2\}} E_Q\left[\left| E_Q[\exp(iuZ_k)|\mathscr{G}_{k-1}] - E_P[\exp(iuX_k)]\right|\right] du \\ & + P[|X_k| > N_k \alpha_k \varrho_k{}^{-1}] + (7N_k + 3) N(0,1)[|\cdot| > \varrho_k] \end{aligned}$$

holds for all $k \in \mathbb{N}$. This follows along the same line as Remark 2.2.

3. Proofs

Let (R, σ_R), (S, σ_S) and (T, σ_T) be Polish spaces. Then the Strassen–Dudley theorem (Dudley (1989), Theorem 11.6.2) can be restated for Markov kernels, which is essential for the proof of the theorem.

Lemma 3.1. *Let K, L be Markov kernels from S to T and let $\beta : S \to [0, \infty[$ be a measurable function. Assume that for a fixed $\alpha > 0$ and for all $x \in S$*

$$\pi(\alpha, K(x, \cdot), L(x, \cdot)) \le \beta(x) .$$

holds. Then for each $\mu \in \mathfrak{M}(S)$ there exists a Markov kernel M from S to $T \times T$ such that for μ-a.e. $x \in S$

$$M(x, \cdot) \in \mathscr{M}(K(x, \cdot), L(x, \cdot)) , \tag{3.1}$$

$$M(x, \{\sigma_T(\cdot, \cdot) > \alpha\}) \le \beta(x) . \tag{3.2}$$

Proof. The existence of a measure $M(x, \cdot)$, $x \in S$ with the desired properties immediately follows by the Strassen-Dudley theorem. All we have to think about is the measurability. For this, we define a multifunction F on S,

$$F(x) := \{\lambda \in \mathfrak{M}(T \times T) | \lambda \in \mathscr{M}(K(x, \cdot), L(x, \cdot)), \lambda(\sigma_T(\cdot, \cdot) > \alpha) \le \beta(x)\} .$$

Denote by $\hat{\mathfrak{B}}(S)$ the (universal) completion of $\mathfrak{B}(S)$. By applying a measurable selection theorem (Sainte-Beuve (1974), Proposition 3) we get a Markov kernel $\hat{M}$ from $(S, \hat{\mathfrak{B}}(S))$ to $(T \times T, \mathfrak{B}(T \times T))$ such that (3.1) and (3.2) hold for all $x \in S$. For $\mu \in \mathfrak{M}(S)$ let $\hat{\mu}$ be the unique extension of μ to $\hat{\mathfrak{B}}(S)$ and define $\hat{\lambda} \in \mathfrak{M}(S \times T \times T, \hat{\mathfrak{B}}(S) \otimes \mathfrak{B}(T \times T))$ by

$$\hat{\lambda}(A) := \int_S \int_{T \times T} 1_A(x, y_1, y_2) \hat{M}(x, dy_1, dy_2) \hat{\mu}(dx) .$$

and consider λ as the restriction of $\hat{\lambda}$ to $\mathfrak{B}(S) \otimes \mathfrak{B}(T \times T)$. Choosing M as a suitable factorization of $\lambda = M \otimes \mu$ we get the appropriate kernel.

Lemma 3.2. *Let $X : (\Omega, \mathfrak{F}, P) \to (S, \mathfrak{B}(S))$ be a random variable, let Q be a Borel law on $S \times T$ with marginal $\mathfrak{L}(X)$ on S. Assume there is a random variable $U : (\Omega, \mathfrak{F}, P) \to (R, \mathfrak{B}(R))$ and a regular version of the conditional distribution $P[U \in du | X = x]$ having no atoms for all $x \in S$. Then there exists a random variable $Y : (\Omega, \mathfrak{F}, P) \to (T, \mathfrak{B}(T))$ with $\mathfrak{L}(X, Y) = Q$.*

This is just a modification of Dudley and Philipp (1983), Lemma 2.11, but in some situations this version is more tractable. The equivalence of both versions immediately can be deduced applying the lemma to a suitable product measure Q. As the proof is nearly the same the reader is refered to Dudley and Philipp (1983).

Proof (of Theorem 1). Lemma 3.1 yields the existence of a Markov kernel M from S to $T \times T$ such that for $Y(Q)$-a.e. $y \in S$

$$M(y, \cdot) \in \mathscr{M}(P[X \in \cdot | Y^* = y], Q[Z \in \cdot | Y = y]) ,$$

$$M(y, \{\sigma_T(\cdot, \cdot) > \alpha\}) \le \beta(y) .$$

Let $K : S \times T \times \mathfrak{B}(R) \to [0, 1]$ be a regular version of $Q[V \in \cdot | (Y, Z) = \cdot]$. We define a law $\lambda \in \mathfrak{M}(S \times T \times T \times R)$ by

$$\lambda(C) := \int\int\int 1_C(y,x,z,v)K(y,z,dv)M(y,dxdz)Y(Q)(dy) .$$

Denoting by π_I the projection corresponding to $I \subset \{1,2,3,4\}$ and remembering $\mathfrak{L}(Y) = \mathfrak{L}(Y^*)$ we get

$$\pi_{\{1,3,4\}}(\lambda) = \mathfrak{L}(Y,Z,V) ,$$

$$\pi_{\{1,2\}}(\lambda) = \mathfrak{L}(Y^*,X) ,$$

$$\pi_{\{2,3\}}(\lambda)(\{\sigma_T(\cdot,\cdot) > \alpha\}) \leq E_Q[\beta(Y)] .$$

In this situation we obtain from Lemma 3.2 random variables Z^*, V^* on $(\Omega, \mathfrak{F}_2, P)$ with values in T and R such that

$$\lambda = \mathfrak{L}(Y^*, X, Z^*, V^*) .$$

By the above mentioned properties of λ the theorem is proved.

Proof (of Remark 2.2). Applying Eberlein (1989), Lemma 2 with $N := N_0\alpha\varrho^{-1}$, $M := 2\varrho^2\alpha^{-1}$, $r := 2^{-1}\alpha$ and $\lambda := N(0, 4^{-1}\alpha^2\varrho^{-2})$ we get for all $y \in \mathbb{R}$

$$\begin{aligned}
&\pi(\alpha, P[X \in \cdot|Y^* = y], Q[Z \in \cdot|Y = y]) \\
&\leq \quad 2\pi^{-1}N_0\alpha\varrho^{-1} \int_{\{|u|\leq 2\alpha^{-1}\varrho^2\}} \\
&\qquad \left|\int_{\mathbb{R}} \exp(iuz)Q[Z \in dz|Y = y] - \int_{\mathbb{R}} \exp(iux)P[X \in dx|Y^* = y]\right| du \\
&\qquad +P[|X| > N_0\alpha\varrho^{-1}|Y^* = y] + (7N_0 + 3)N(0,1)[|\cdot| > \varrho] .
\end{aligned}$$

We define $\beta(y)$ as the right-hand side of this inequality. Because of the relation between regular conditional distribution and conditional expectation we get

$$\begin{aligned}
\beta(Y) \quad = \quad & 2\pi^{-1}N_0\alpha\varrho^{-1} \int_{\{|u|\leq 2\alpha^{-1}\varrho^2\}} \\
& \left|E_Q[\exp(iuZ)|\sigma(Y)] - E_P[\exp(iuX)|Y^* = \cdot](Y)\right| du \\
& +P[|X| > N_0\alpha\varrho^{-1}|Y^* = \cdot](Y) + (7N_0 + 3)N(0,1)[|\cdot| > \varrho] .
\end{aligned}$$

and hence the remark is proved.

Proof (of Remark 2.3). Recalling the notation of Theorem 1 in Monrad and Philipp (1991) let X be a $\mathbb{R}$-valued random variable defined on a probability space $(\Omega, \mathfrak{F}, P)$ and let $\mathfrak{F}_1$ be a countably generated sub-σ-field of $\mathfrak{F}$. Let U be a uniformly distributed random variable on $(\Omega, \mathfrak{F}, P)$ assuming values in $[0,1]$ that is independent of the σ-field generated by $\mathfrak{F}_1$ and $\sigma(X)$. Let $G[\cdot|\mathfrak{F}_1]$ be a regular conditional distribution on $\mathbb{R}$ with conditional characteristic function $g(\cdot|\mathfrak{F}_1)$. Suppose that for some numbers $\lambda, \delta > 0$ and $T \geq 2$

$$\int_{\{|u|\le T\}} E\left[\left|E[\exp(iuX)|\mathfrak{F}_1] - g(u|\mathfrak{F}_1)\right|\right] du \le 2\lambda T$$

and

$$E[G[|\cdot| \ge 2^{-1}T|\mathfrak{F}_1]] < \delta\ .$$

Since the sub-σ-field $\mathfrak{F}_1$ is countably generated there exists a real valued random variable Y such that $\sigma(Y) = \mathfrak{F}_1$. According to Monrad and Philipp (1991), Lemma 2.2.1 we find a $\mathbb{R}$-valued random variable Z on $(\Omega, \mathfrak{F}, P)$ and a regular version of the conditional distribution with $P[Z \in \cdot|Y = \cdot](Y) = G[\cdot|\mathfrak{F}_1](\cdot)$. For the application of Theorem 1 the random variable V is negligible, so we take $V := 1$. Denoting $Y^* = Y$ Theorem 1 together with Remark 2.2 yields the existence of a $\mathbb{R}$-valued random variable Z^* on $(\Omega, \mathfrak{F}, P)$ such that $\mathfrak{L}(Y, Z) = \mathfrak{L}(Y^*, Z^*)$ and

$$P[|X - Z^*| > \alpha] \le 4\pi^{-1}N_0\alpha\varrho^{-1}\lambda T + \delta + (3 + 7N_0)N(0,1)(|\cdot| > \varrho)\ .$$

With $N_0 := \alpha(2^{-1}T\alpha^{-1})^{3/2}$ and $\varrho^2 := 2^{-1}T\alpha$ we obtain the statement of the remark.

Acknowledgement. This work is part of the author's PhD thesis. It was supported by the Deutsche Forschungsgemeinschaft under grant Eb 66/3-3. I want to thank Professor Dr. E. Eberlein for valuable suggestions and for his interest in this work.

References

Berkes, I., Philipp, W.: Approximation theorems for independent and weakly dependent random vectors. Ann. Probab. **7**, 29–54 (1979)

Besdziek, N.: Strong approximations of semimartingales by processes with independent increments. Probab. Theory Relat. Fields **87**, 489–520 (1991)

Dudley, R.M.: Real Analysis and Probability. Belmont: Wadsworth and Brooks/Cole 1989

Dudley, R.M., Philipp, W.: Invariance principles for sums of Banach space valued random elements and empirical processes. Z. Wahrscheinlichkeitstheor. Verw. Geb. **62**, 509–552 (1983)

Eberlein, E.: Strong approximation of continuous time stochastic processes. J. Multivariate Anal. **31**, 220–235 (1989)

Eberlein, E.: On modelling questions in security valuation. Math. Finance **2**, 17–32 (1992)

Eberlein, E., Römersperger, M.: Strong approximation of semimartingales and statistical processes. Probab. Theory Relat. Fields **104**, 539–567 (1996)

Monrad, D., Philipp, W.: Nearby variables with nearby conditional laws and a strong approximation theorem for Hilbert space valued Martingales. Probab. Theory Relat. Fields **88**, 381–404 (1991)

Monrad, D., Philipp, W.: Conditional versions of the Strassen-Dudley theorem. In: Dudley et al. (eds.). Probability in Banach spaces 8, pp. 116–128. Boston Basel: Birkhäuser 1992

Morrow, G., Philipp, W.: An almost sure invariance principle for Hilbert space valued martingales. Trans. Am. Math. Soc. **273**, 231–251 (1982)

Philipp, W.: A note on the almost sure approximation of weakly dependent random variables. Monatsh. Math. **102**, 227–236 (1986)

Römersperger, M.: Starke Approximationen für Semimartingale und stochastische Prozesse in der Statistik. Dissertation Universität Freiburg, 1993

Sainte-Beuve, M.-F.: On the extension of von Neumann-Aumann's theorem. J. Functional Anal. **17**, 112–129 (1974)

Strassen, V.: An Invariance Principle for the Law of the Iterated Logarithm. Z. Wahrscheinlichkeitstheor. Verw. Geb. **3**, 211–226 (1964)

Strittmatter, W.: Measures of dependence for processes in metric spaces. Stochastics **27**, 33–50 (1989)

This article was processed by the author using the LaTeX style file *pljour1m* from Springer-Verlag.