The Conditional Strong Law of Large Numbers in Separable Banach Spaces

Takuhisa Shikimi

Abstract

Let $\mathscr{X}$ be a separable Banach space. Suppose that $X_1, X_2, \ldots$ are $\mathscr{X}$-valued random variables that are conditionally independent and identically distributed given a sub-σ-algebra. We show that, conditional on the sub-σ-algebra, $n^{-1}\Sigma_{i=1}^{n} X_i$ converges to the conditional expectation of X_1 a.s.

Keywords : conditional strong law of large numbers; separable Banach spaces; exchangeability

1 Introduction and Preliminaries

Our purpose in this article is to prove the conditional version of the Kolmogorov's strong law of large numbers (SLLN) for conditionally independent and identically distributed random variables, each one defined on a probability space and taking values in a separable Banach space. Majerek et al. (2005) showed conditional versions of the Borel-Cantelli lemma and Kolmogorov's maximal inequality and proved the conditional SLLN for real-valued random variables. Prakasa Rao (2009) derived a conditional version of the generalized Kolmogorov's maximal inequality due to Loève (1977, p.275) to obtain the conditional SLLN for real-valued random variables.

It is well known that the SLLN extends to random variables taking values in a separable Banach space. If $X_1, X_2,...$ are i.i.d. random variables with values in a separable Banach space $\mathscr{X}$ and X_1 is integrable, then $S_n/n \to \mathbf{E}X_1$, where $S_n = X_1 + \cdots + X_n$. As in the case of real-valued random variables, it is natural to expect that if $\mathscr{G}$ is a sub-σ-algebra of $\mathscr{F}$ and $X_1, X_2,...$ are conditionally independent and identically distributed given $\mathscr{G}$, then the conditional version of the SLLN holds:

$$\mathbf{P}\left(\frac{S_n}{n} \to \mathbf{E}(X_1 \mid \mathscr{G}) \mid \mathscr{G}\right) = 1 \text{ a.s.} \tag{1.1}$$

We can prove (1.1) by utilizing the approach due to Chow and Teicher (1997, p.235) together with the SLLN for i.i.d. random variable with values in a separable Banach space. Etemadi (1997) proved the SLLN for 2-exchangeable (and hence exchangeable) integrable random variables in a separable Banach space. Taking expectations on both sides of (1.1) gives an alternative proof of the SLLN for exchangeable random variables in a separable Banach space, since a sequence of random variables is exchangeable if and only if it is conditionally i.i.d. given some sub-σ-algebra.

Unless otherwise stated, we always assume that $\mathscr{X}$ is a separable Banach space with norm $\|\cdot\|$. Let $\mathscr{X}$ be equipped with the Borel σ-algebra $\mathscr{A}$, which is the σ-algebra generated by the norm topology τ. The space of all bounded linear functionals on $\mathscr{X}$, denoted by $\mathscr{X}^*$, is called the *dual* of $\mathscr{X}$. For $x \in \mathscr{X}$ and $x^* \in \mathscr{X}^*$, $x^*(x)$ is denoted by $\langle x,x^* \rangle$. The dual $\mathscr{X}^*$ is a Banach space with norm $\|x^*\| = \sup_{\|x\| \leqslant 1} |\langle x,x^* \rangle|$. Further, since $\|x\| = \sup_{\|x^*\| \leqslant 1} |\langle x,x^* \rangle|$ for every $x \in \mathscr{X}$, the dual $\mathscr{X}^*$ separates points in

$\mathscr{X}$. Namely, if $\langle x,x^*\rangle=0$ for all $x^*\in\mathscr{X}^*$, then $x=0$. Since $\mathscr{X}$ is separable, $\mathscr{X}^*$ contains a countable set $\{x_n^*\}$ that separates points in $\mathscr{X}$.

The product topological space $(\mathscr{X}^\infty,\tau^\infty):=(\mathscr{X}\times\mathscr{X}\times\cdots,\ \tau\times\tau\times\cdots)$ is a Polish space with the metric $\rho(x,y):=\Sigma_{i=1}^{\infty}2^{-i}\|x_i-y_i\|/(1+\|x_i-y_i\|)$, where $x=(x_i)$ and $y=(y_i)$ are elements in $\mathscr{X}^\infty$. The product σ-algebra $\mathscr{A}^\infty:=\mathscr{A}\times\mathscr{A}\times\cdots$ is generated by τ^∞ since $\mathscr{X}$ is separable.

Let $(\Omega,\mathscr{F},\mathbf{P})$ be a probability space. A random variable $X:\Omega\to\mathscr{X}$ is called *integrable* if $\mathbf{E}\|X\|<\infty$. If X is integrable, $\langle X,x^*\rangle$ is integrable for every $x^*\in\mathscr{X}^*$ since $|\langle X,x^*\rangle|\leqslant\|X\|\|x^*\|$, and there exists an element $m\in\mathscr{X}$ such that $\langle m,x^*\rangle=\mathbf{E}\langle X,x^*\rangle$ for every $x^*\in\mathscr{X}^*$. It is clear that there is at most one such m because $\mathscr{X}^*$ separates points in $\mathscr{X}$. Such m is called the *expectation* of X and denoted by $\mathbf{E}X$ or $\int Xd\mathbf{P}$. If μ is the distribution of X, $\mathbf{E}X=\int x\mu(dx)$. If X is integrable, so is 1_HX, and the integral of X over H is defined by $\int_H Xd\mathbf{P}=\mathbf{E}(1_HX)$.

Given a sub-σ-algebra $\mathscr{G}$ of $\mathscr{F}$, an integrable random variable Y with values in $\mathscr{X}$ is called the *conditional expectation* of X given $\mathscr{G}$ if Y is $\mathscr{G}$-measurable and

$$\int_G Yd\mathbf{P}=\int_G Xd\mathbf{P}$$

for all $G\in\mathscr{G}$. For the existence of the conditional expectation, we refer to Stroock (1993, Theorem 5.1.22). Any version of the conditional expectation of X given $\mathscr{G}$ is denoted by $\mathbf{E}(X|\mathscr{G})$. For every $x^*\in\mathscr{X}^*$, $\langle\mathbf{E}(X|\mathscr{G}),x^*\rangle$ is clearly $\mathscr{G}$-measurable. It is integrable because $|\langle\mathbf{E}(X|\mathscr{G}),x^*\rangle|\leqslant\|\mathbf{E}(X|\mathscr{G})\|\|x^*\|$. Given $G\in\mathscr{G}$,

$$\int_G\langle\mathbf{E}(X|\mathscr{G}),x^*\rangle d\mathbf{P}=\left\langle\int_G\mathbf{E}(X|\mathscr{G})d\mathbf{P},x^*\right\rangle=\left\langle\int_G Xd\mathbf{P},x^*\right\rangle$$

$$= \int_G \langle X,x^* \rangle d\mathbf{P}.$$

Hence, $\langle \mathbf{E}(X|\mathscr{G}), x^* \rangle$ is a version of the conditional expectation of $\langle X,x^* \rangle$ given $\mathscr{G}$ for every $x^* \in \mathscr{X}^*$.

We say that $X_1,X_2,\cdots$ are *conditionally independent* given $\mathscr{G}$ if for all $n \geqslant 1$ and all $A_1,\cdots,A_n \in \mathscr{A}$,

$$\mathbf{P}(X_1 \in A_1,\cdots,Xn \in A_n|\mathscr{G}) = \mathbf{P}(X_1 \in A_1|\mathscr{G})\cdots\mathbf{P}(X_1 \in A_1|\mathscr{G}) \text{ a.s.}$$

We say that $X_1,X_2,\cdots$ are *conditionally identically distributed* if for all $n \geqslant 1$ and all $A \in \mathscr{A}$,

$$\mathbf{P}(X_n \in A\,|\mathscr{G}) = \mathbf{P}(X_1 \in A\,|\mathscr{G}) \text{ a.s.}$$

If $X_1,X_2,\cdots$ are conditionally independent and identically distributed given $\mathscr{G}$, they are called *conditionally i.i.d.* given $\mathscr{G}$ for short.

The proof of our main theorem (Theorem 2.2) relies on the following lemmas, whose proof are in Section 3.

Lemma 1.1. *Let* $(\mathscr{X}_i, \mathscr{A}_i)$ *be a measurable space and let* X_i *be a* $\mathscr{X}_i$*- valued random variable* $(i=1,2)$. *Suppose that* X_2 *is* $\mathscr{G}$*-measurable and that there exists a regular conditional distribution* $(\omega,A_1) \in \Omega \times \mathscr{A}_1 \to \mu^\omega(A_1) \in [0,1]$ *for* X_1 *given* $\mathscr{G}$. *If* $f: \mathscr{X}_1 \times \mathscr{X}_2 \to \mathscr{X}$ *is a measurable function such that* $f(X_1,X_2)$ *is integrable, then* **P**-*almost all* $\omega \in \Omega$, $f(\cdot,X_2(\omega))$ *is integrable with respect to* $\mu^\omega(\cdot)$ *and*

$$\mathbf{E}(f(X_1,X_2)\,|\mathscr{G})(\omega) = \int f(x_1,X_2(\omega))\mu^\omega(dx_1). \tag{1.2}$$

For a sequence $X_1,X_2,...$ of $\mathscr{X}$-valued random variables, $X=(X_1,X_2,...)$ has a regular conditional distribution given $\mathscr{G}$ because $(\mathscr{X}^\infty,\mathscr{A}^\infty)$ is a

Borel space.

Lemma 1.2. *Suppose that* $X_1, X_2, \ldots$ *are* $\mathscr{X}$*-valued random variables that are conditionally i.i.d. given* $\mathscr{G}$*, and for each* $(\omega, A) \in \Omega \times \mathscr{A}^\infty$ *let* $\mu^\omega(A)$ *be a regular conditional distribution for* $X = (X_1, X_2, \ldots)$ *given* $\mathscr{G}$*. Then, there exists a* **P***-null set* N *in* $\mathscr{G}$ *such that the following properties hold:*

(i) *for all* $\omega \notin N$*, the coordinate functions* $\xi_1, \xi_2, \ldots$ *on* $(\mathscr{X}^\infty, \mathscr{A}^\infty, \mu^\omega(\cdot))$ *are i.i.d. ;*

(ii) *if* X_1 *is integrable, then for all* $\omega \notin N$*,* $\xi_1, \xi_2, \ldots$ *are integrable with respect to* $\mu^\omega(\cdot)$

2 The Conditional SLLN

The following lemma is a simple generalization of a well-known convergence criterion. We omit the proof because it is proved in much the same way as the unconditional version.

Lemma 2.1. *Let* $(\mathscr{X}, \|\cdot\|)$ *be a separable normed space. Suppose that* Z*,* $Z_1, Z_2, \ldots$ *are* $\mathscr{X}$*-valued random variables defined on* $(\Omega, \mathscr{F}, \mathbf{P})$*. Then, the following assertions are equivalent :*

(i) $\mathbf{P}(Z_n \to Z \mid \mathscr{G}) = 1$ *a.s. ;*

(ii) $\forall \varepsilon > 0 \ \lim_n \mathbf{P}(\sup_{k \geqslant n} \|Z_k - Z\| > \varepsilon \mid \mathscr{G}) = 0$ *a.s.*

Theorem 2.2 (**Conditional strong law of large numbers**). *Suppose that* $X_1, X_2, \ldots$ *are* $\mathscr{X}$*-valued integrable random variables that are conditionally i.i.d. given* $\mathscr{G}$*, and let* $S_n = X_1 + \cdots + X_n, n \geqslant 1$*. Then,*

$$\mathbf{P}\left(\frac{S_n}{n} \to \mathbf{E}(X_1 \mid \mathscr{G}) \mid \mathscr{G}\right) = 1 \text{ a.s.} \tag{2.1}$$

If

$$\mathbf{P}\left(\frac{S_n}{n} \to m \mid \mathscr{G}\right) = 1 \text{ a.s.} \tag{2.2}$$

then $m = \mathbf{E}(X_1 \mid \mathscr{G})$ *a.s.*

Proof. We prove that

$$\forall \varepsilon > 0 \lim_n \mathbf{P}\left(\sup_{k \geqslant n} \left\| \frac{S_k}{k} - \mathbf{E}(X_1 \mid \mathscr{G}) \right\| > \varepsilon \,\middle|\, \mathscr{G}\right) = 0 \text{ a.s.} \tag{2.2}$$

which is equivalent to (2.1) by Lemma 2.1. Let $\mu^\omega(A)$ be a regular conditional distribution for $X = (X_1, X_2, \ldots)$ given $\mathscr{G}$. Let $\xi_1, \xi_2, \ldots$ be a sequence of coordinate functions on $(\mathscr{X}^\infty, \mathscr{A}^\infty)$ as in Lemma 1.2. Since $\xi_1(X)$ (= X_1) is integrable, Lemma 1.1 shows that

$$\begin{aligned}\mathbf{E}(X_1 \mid \mathscr{G})(\omega) &= \mathbf{E}(\xi_1(X) \mid \mathscr{G})(\omega) \\ &= \int \xi_1(x)\mu^\omega(dx)\ \mathbf{P}\text{-almost all } \omega \in \Omega.\end{aligned}$$

Fix $\varepsilon > 0$. Since $\mathbf{E}(X_1 \mid \mathscr{G})$ is $\mathscr{G}$-measurable, Lemma 1.1 ensures that for $\mathbf{P}$-almost all $\omega \in \Omega$,

$$\begin{aligned}&\mathbf{P}\left(\sup_{k \geqslant n} \left\| \frac{S_k}{k} - \mathbf{E}(X_1 \mid \mathscr{G}) \right\| > \varepsilon \,\middle|\, \mathscr{G}\right)(\omega) \\ &= \mathbf{P}\left(\sup_{k \geqslant n} \left\| \frac{1}{k}\sum_{i=1}^{k} \xi_1(X) - \mathbf{E}(X_1 \mid \mathscr{G}) \right\| > \varepsilon \,\middle|\, \mathscr{G}\right)(\omega) \\ &= \mu^\omega\left(x \in \mathscr{X}^\infty : \sup_{k \geqslant n} \left\| \frac{1}{k}\sum_{i=1}^{k} \xi_i(x) - \mathbf{E}(X_1 \mid \mathscr{G})(\omega) \right\| > \varepsilon\right) \qquad (2.4) \\ &= \mu^\omega\left(x \in \mathscr{X}^\infty : \sup_{k \geqslant n} \left\| \frac{1}{k}\sum_{i=1}^{k} \xi_i(x) - \int \xi_1(x)\mu^\omega(dx) \right\| > \varepsilon\right).\end{aligned}$$

By Lemma 1.2, for $\mathbf{P}$-almost all $\omega \in \Omega$, $\xi_1, \xi_2, \ldots$ are i.i.d. integrable random variables on $(\mathscr{X}^\infty, \mathscr{A}^\infty, \mu^\omega(\cdot))$. It follows from the SLLN for $\mathscr{X}$-valued random variables (Stroock (1993, pp.137-139)) that

$$\mu^{\omega}\left(x\in\mathscr{X}^{\infty}:\sup_{k\geqslant n}\left\|\frac{1}{k}\sum_{i=1}^{k}\xi_i(x)-\int\xi_1(x)\,d\mu^{\omega}(dx)\right\|>\varepsilon\right)\to 0$$

for **P**-almost all $\omega\in\Omega$, which, together with (2.4), yields (2.3). If (2.2) holds, $m=\mathbf{E}(X_1\,|\mathscr{G})$ a.s. by the result just proved above. □

A sequence $X_1,X_2,\ldots$ of $\mathscr{X}$-valued random variables is said to be *exchangeable* if the distribution of $(X_{\sigma(1)},\ldots,X_{\sigma(n)})$ is invariant for each $n\geqslant 1$ and each permutation σ of $\{1,\ldots,n\}$. According to de Finetti's theorem, $X_1,X_2,\ldots$ is exchangeable if and only if the random variables are conditionally i.i.d. given some sub-σ- algebra $\mathscr{G}$. We can take $\mathscr{G}$ to be the tail σ-algebra of $X_1,X_2,\ldots$. As a corollary of Theorem 2.2 we see that if $X_1, X_2,\ldots$ is an exchangeable sequence of $\mathscr{X}$-valued integrable random variables, then $S_n/n\to\mathbf{E}(X_1\,|\mathscr{G})$ for some sub-σ-algebral $\mathscr{G}$.

3 Proofs

Proof of Lemma 1.1. By the integrability condition, $\mathbf{E}(\|f(X_1,X_2)\|\,|\mathscr{G})<\infty$ a.s. Since

$$\mathbf{E}(\|f(X_1,X_2)\|\,|\mathscr{G})=\int\|f(x_1,X_2(\omega))\|\,\mu^{\omega}(dx_1)\text{ a.s.,}$$

$f(\cdot,X_2(\omega))$ is integrable with respect to $\mu^{\omega}(\cdot)$ for **P**-almost all $\omega\in\Omega$. Let N be a **P**-null set on which $f(\cdot,X_2(\omega))$ fails to be integrable with respect to $\mu^{\omega}(\cdot)$. Let $D\subset\mathscr{X}^*$ be a countable set that separates points in $\mathscr{X}$. Fix $x^*\in D$. For $\omega\notin N$, $\langle f(\cdot,X_2(\omega),x^*)\rangle$ is integrable with respect to $\mu^{\omega}(\cdot)$ and

$$\left\langle\int f(x_1,X_2(\omega))\,\mu^{\omega}(dx_1),x^*\right\rangle=\int\langle f(x_1,X_2(\omega)),x^*\rangle\,\mu^{\omega}(dx_1).$$

On the other hand,

$$\langle \mathbf{E}(f(X_1,X_2)\,|\mathscr{G})(\omega),x^*\rangle = \mathbf{E}(\langle f(X_1,X_2),x^*\rangle\,|\mathscr{G})(\omega) \text{ a.s.}$$
$$= \int\langle f(x_1,X_2(\omega)),x^*\rangle\,\mu^\omega(dx) \text{ a.s.} \qquad (3.1)$$

Thus, there exists a $\mathbf{P}$-null set $N(x^*)$ such that for $\omega\notin N(x^*)$, we have (3.1). It follows that for all $\omega\notin N\cup_{x^*\in D}N(x^*)$

$$\langle \mathbf{E}(f(X_1,X_2)\,|\mathscr{G})(\omega),x^*\rangle = \left\langle \int f(x_1,X_2(\omega))\,\mu^\omega(dx_1),x^*\right\rangle$$

for every $x^*\in D$. Hence, for all $\omega\notin N\cup_{x^*\in D}N(x^*)$, $f(\cdot,X_2(\omega))$ is integrable with respect to μ^ω, and (1.2) holds. □

Proof of Lemma 1.2. Let $\mathscr{B}$ denote a countable base for τ and $\mathscr{B}'$ denote the class of finite intersections of sets in $\mathscr{B}$. Notice that the class

$$\mathscr{J} = \{B_1\times\cdots\times B_n\times\mathscr{X}\times\cdots : n\geqslant 1, B_1\times\cdots\times B_n\in\mathscr{B}'\}$$

is a countable π-class that generates $\mathscr{A}^\infty$.

(i) For each $i\geqslant 1$ and $B\in\mathscr{B}'$,

$$\mu^\omega(\xi_i\in B) = \mu^\omega(\mathscr{X}\times\cdots\times\mathscr{X}\times B\times\mathscr{X}\times\cdots)$$
$$= \mathbf{P}(X\in\mathscr{X}\times\cdots\times\mathscr{X}\times B\times\mathscr{X}\times\cdots\,|\mathscr{G})(\omega) \text{ a.s.}$$
$$= \mathbf{P}(X_i\in B\,|\mathscr{G})(\omega)$$
$$= \mathbf{P}(X_1\in B\,|\mathscr{G})(\omega) \text{ a.s.}$$
$$= \mu^\omega(\xi_1\in B) \text{ a.s.}$$

Let $M_{i,B}$ be a $\mathbf{P}$-null set such that $\mu^\omega(\xi_i\in B) = \mu^\omega(\xi_1\in B)$ outside $M_{i,B}$. The union $\cup_{i\geqslant 1,B\in\mathscr{B}'} M_{i,B}$ belongs to $\mathscr{G}$ since every $M_{i,B}$ lies in $\mathscr{G}$. If $\omega\notin M := \cup_{i\geqslant 1,B\in\mathscr{B}'}$, then $\mu^\omega(\xi_i\in B) = \mu^\omega(\xi_1\in B)$ for all $i\geqslant 1$ and $B\in\mathscr{B}'$. Since $\mathscr{B}'$ is a π-class generating $\mathscr{A}$, for all $\omega\notin M$,

$$\mu^{\omega}(\xi_1\in\cdot)=\mu^{\omega}(\xi_2\in\cdot)=\cdots.$$

For each $\omega\in\Omega$, define the product probability measure $\nu^{\omega}(\cdot)$ on $(\mathscr{X}^{\infty}, \mathscr{A}^{\infty})$ by

$$\nu^{\omega}(\cdot)=\mu^{\omega}(\xi_1\in\cdot)\times\mu^{\omega}(\xi_2\in\cdot)\times\cdots.$$

For all $n\geqslant 1$ and $B_1,\dots,B_n\in\mathscr{B}'$,

$$\begin{aligned}\mu^{\omega}(B_1\times\cdots\times B_n\times\mathscr{X}\times\cdots)&=\mathbf{P}(X_1\in B_1,\dots,X_n\in B_n\mid\mathscr{G})(\omega)\ \text{a.s.}\\&=\mathbf{P}(X_1\in B_1\mid\mathscr{G})(\omega)\cdots\mathbf{P}(X_n\in B_n\mid\mathscr{G})(\omega)\text{a.s.}\\&=\mu^{\omega}(\xi_1\in B_1)\cdots\mu^{\omega}(\xi_n\in B_n)\ \text{a.s.}\\&=\nu^{\omega}(B_1\times\cdots\times B_n\times\mathscr{X}\times\cdots).\end{aligned}$$

Let $M(B_1,\dots,B_n)$ be a $\mathbf{P}$-null set such that $\mu^{\omega}(B_1\times\cdots\times B_n\times\mathscr{X}\times\cdots)=\nu^{\omega}(B_1\times\cdots\times B_n\times\mathscr{X}\times\cdots)$ for all $\omega\notin M(B_1,\dots,B_n)$, and let $M_n=\cup_{B_1,\dots,B_n\in\mathscr{B}'}M(B_1,\dots,B_n)$. Then, $\omega\notin M':=\cup_{n\geqslant1}M_n$, $\mu^{\omega}(B_1\times\cdots\times B_n\times\mathscr{X}\times\cdots)=\nu^{\omega}(B_1\times\cdots\times B_n\times\mathscr{X}\times\cdots)$ for all $n\geqslant1$ and $B_1,\dots,B_n\in\mathscr{B}'$. Since $\mathscr{J}$ is a π-class that generates $\mathscr{A}^{\infty}$, we have $\mu^{\omega}=\nu^{\omega}$ for all $\omega\notin M'$. In particular, for all $\omega\notin M'$, $n\geqslant1$ and $A_1,\dots,A_n\in\mathscr{A}$,

$$\begin{aligned}\mu^{\omega}(\xi_1\in A_1,\dots,\xi_n\in A_n)&=\mu^{\omega}(A_1\times\cdots\times A_n\times\mathscr{X}\times\cdots)\\&=\nu^{\omega}(A_1\times\cdots\times A_n\times\mathscr{X}\times\cdots)\\&=\mu^{\omega}(\xi_1\in A_1)\cdots\mu^{\omega}(\xi_n\in A_n).\end{aligned}$$

As a result, for all $\omega\notin M\cup M'$, $\xi_1,\xi_2,\dots,$ are i.i.d. under $\mu^{\omega}(\cdot)$.

(ii) Assume that X_1 is integrable. Since $\mathbf{E}(\|\xi_1(X)\|\mid\mathscr{G})=\mathbf{E}(\|X_1\|\mid\mathscr{G})<\infty$ a.s. and

$$\mathbf{E}(\|\xi_1(X)\|\mid\mathscr{G})(\omega)=\int\|\xi_1(x)\|\,\mu^{\omega}(dx)\,\mathbf{P}\text{-a.s.},$$

there exists a **P**-null set $M''\in\mathscr{G}$ such that for all $\omega\notin M''$,

$$\int \|\xi_1(x)\|\,\mu^\omega(dx)<\infty.$$

Putting $N=M\cup M'\cup N'$ completes the proof. □

References

Chow, Y.S. and Teicher, H., 1997. Probability Theory, 3rd ed. Springer, New York.

Etemadi, N., 1997. Criteria for the strong law of large numbers for sequences of arbitrary random vectors. Statist. Probab. Lett. **33**, 151–157.

Loève, M., 1977. Probability Theory I, 4th ed. Springer, New York.

Majerek, D., Nowak, W., Zieba, W., 2005. Conditional strong law of large number. Int. J. Pure Appl. Math. **20**(2), 143–157.

Prakasa Rao, B.L.S., 2009. Conditional independence, conditional mixing and conditional association. Ann. Inst. Stat. Math. **61**, 441–460.

Stroock, D.W., 1993. Probability Theory: An Analytic View, revised ed. Cambridge University Press, Cambridge.