Exchangeability and de Finetti's theorem

Cameron Freer

March 3, 2026

Chapter 1

Introduction

This blueprint documents the formalization of **de Finetti's theorem** and the **de Finetti–Ryll-Nardzewski equivalence** for infinite sequences on *standard Borel spaces*.

The main result establishes a three-way equivalence between:

- **Contractable**: All strictly increasing subsequences of equal length have the same distribution
- **Exchangeable**: Distribution invariant under finite permutations
- **Conditionally i.i.d.**: There exists a probability kernel such that finite marginals equal mixtures of product measures

We formalize *all three proofs* from Kallenberg (2005):

1. **Koopman/Ergodic approach** using the Mean Ergodic Theorem
2. **L^2 approach** using elementary contractability bounds
3. **Martingale approach** using reverse martingale convergence (after Aldous)

Chapter 2

Foundations

2.1 Core Definitions

Definition 1 (Exchangeable sequence)**.** A sequence $(X_n)_{n\in\mathbb{N}}$ of random variables is *exchangeable* with respect to measure μ if for every $n \in \mathbb{N}$ and every permutation σ of $\{0, \dots, n-1\}$, the joint distribution of $(X_{\sigma(0)}, \dots, X_{\sigma(n-1)})$ equals that of $(X_0, \dots, X_{n-1})$.

Definition 2 (Contractable sequence)**.** A sequence $(X_n)_{n\in\mathbb{N}}$ is *contractable* with respect to measure μ if for all $m \in \mathbb{N}$ and all strictly increasing functions $k, k' : \mathrm{Fin}(m) \to \mathbb{N}$, the joint distribution of $(X_{k(0)}, \dots, X_{k(m-1)})$ equals that of $(X_{k'(0)}, \dots, X_{k'(m-1)})$.

Definition 3 (Conditionally i.i.d. sequence)**.** A sequence $(X_n)_{n\in\mathbb{N}}$ is *conditionally i.i.d.* with respect to measure μ if there exists a probability kernel $\nu : \Omega \to \mathrm{Measure}(\alpha)$ such that for any strictly monotone $k : \mathrm{Fin}(m) \to \mathbb{N}$, the joint distribution of $(X_{k(0)}, \dots, X_{k(m-1)})$ equals the mixture $\mu.\mathrm{bind}(\omega \mapsto \nu(\omega)^{\otimes m})$.

2.2 σ-algebra Infrastructure

Definition 4 (Tail σ-algebra)**.** The *tail σ-algebra* of a sequence (X_n) is $\mathcal{T} = \bigcap_{n=0}^{\infty} \sigma(X_n, X_{n+1}, \dots)$.

Definition 5 (Future filtration)**.** The *future filtration* at level m is $\mathcal{F}_m = \sigma(X_{m+1}, X_{m+2}, \dots)$.

Lemma 6 (Future filtration is antitone)**.** *The future filtration is antitone: $m \le n$ implies $\mathcal{F}_n \le \mathcal{F}_m$.*

Lemma 7 (Tail σ-algebra contained in future filtration)**.** *For all m: $\mathcal{T} \le \mathcal{F}_m$.*

Lemma 8 (Tail σ-algebra is infimum of reverse filtration)**.** $\mathcal{T} = \bigwedge_{m=0}^{\infty} \mathcal{F}_m$.

Definition 9 (Shift-invariant σ-algebra)**.** The shift-invariant σ-algebra consists of sets S such that $\theta^{-1}(S) = S$ where θ is the shift operator.

Chapter 3

Easy Directions

3.1 Exchangeable implies Contractable

Lemma 10 (Permutation extension)**.** *Any strictly increasing function* $k : \mathrm{Fin}(m) \to \mathbb{N}$ *with range contained in* $\{0, \dots, n-1\}$ *extends to a permutation of* $\{0, \dots, n-1\}$*.*

Theorem 11 (Exchangeable implies Contractable)**.** *If* (X_n) *is exchangeable, then it is contractable.*

3.2 Conditionally i.i.d. implies Exchangeable

Theorem 12 (Conditionally i.i.d. implies Exchangeable)**.** *If* (X_n) *is conditionally i.i.d., then it is exchangeable.*

Chapter 4

Main Implication: Contractable implies Conditionally i.i.d.

This is the deep direction of de Finetti's theorem. We formalize three independent proofs.

4.1 Via Martingale (Aldous' proof)

The martingale approach uses reverse martingale convergence to the tail σ-algebra.

4.1.1 Pair Law Equality

Lemma 13 (Pair law equality for contractable sequences)**.** *For a contractable sequence and $k \leq m$, the joint distribution of $(X_k, X_{m+1}, X_{m+2}, \ldots)$ equals that of $(X_m, X_{m+1}, X_{m+2}, \ldots)$.*

Lemma 14 (Contractable distribution equality)**.** *For a contractable sequence, the joint distribution of $(X_k, \theta_{m+1} X)$ equals that of $(X_m, \theta_{m+1} X)$ for all $k \leq m$, where θ_n is the shift operator.*

Lemma 15 (Conditional expectation of indicator equals under contractability)**.** *For contractable sequences and $k \leq m$: $\mathbb{E}[\mathbf{1}_{X_k \in B} \mid \mathcal{F}_m] = \mathbb{E}[\mathbf{1}_{X_m \in B} \mid \mathcal{F}_m]$ a.s.*

4.1.2 Kallenberg Chain and Convergence

Lemma 16 (Kallenberg chain lemma)**.** *Conditional expectations of indicators given the reverse filtration converge to conditional expectations given the tail σ-algebra.*

Lemma 17 (Conditional expectation convergence)**.** *For $k \leq m$ and measurable B: $\mathbb{E}[\mathbf{1}_{X_m \in B} \mid \mathcal{F}_m] = \mathbb{E}[\mathbf{1}_{X_k \in B} \mid \mathcal{F}_m]$ a.s.*

Lemma 18 (Extreme members equal on tail)**.** *For any measurable B: $\mathbb{E}[\mathbf{1}_{X_m \in B} \mid \mathcal{T}] = \mathbb{E}[\mathbf{1}_{X_0 \in B} \mid \mathcal{T}]$ a.s.*

4.1.3 Factorization and Directing Measure

Lemma 19 (Finite level factorization)**.** *For finite products of indicators, the conditional expectation given $\mathcal{F}_m$ factors as a product of individual conditional expectations.*

Lemma 20 (Tail factorization from future)**.** *The tail σ-algebra factorization follows from the finite-level factorization via reverse martingale convergence.*

Lemma 21 (Block coordinate conditional independence)**.** *Coordinates are conditionally independent given the tail σ-algebra.*

Lemma 22 (Directing measure is probability measure)**.** *The directing measure $\nu(\omega)(B) = \mathbb{E}[\mathbf{1}_{X_0 \in B} \mid \mathcal{T}](\omega)$ is a probability measure for a.e. ω.*

Lemma 23 (Directing measure measurability)**.** *The directing measure $\omega \mapsto \nu(\omega)(B)$ is measurable for each Borel B.*

Theorem 24 (Contractable implies Conditionally i.i.d. (via Martingale))**.** *If (X_n) is contractable, then it is conditionally i.i.d. The directing kernel $\nu(\omega)(B) = \mathbb{E}[\mathbf{1}_{X_0 \in B} \mid \mathcal{T}](\omega)$ is constructed from the tail σ-algebra.*

4.2 Via L^2 (Elementary proof)

The L^2 approach uses elementary contractability bounds on block averages. This is Kallenberg's "second proof" and has the lightest dependencies.

Note: This proof applies to *real-valued* sequences ($X : \mathbb{N} \to \Omega \to \mathbb{R}$) with L^2 integrability (i.e., $\mathbb{E}[X_i^2] < \infty$ for all i).

4.2.1 Block Averages and Covariance Structure

Definition 25 (Block average (L^2 version))**.** The block average $A_n = \frac{1}{n} \sum_{i=0}^{n-1} f(X_i)$ for bounded measurable f.

Lemma 26 (Contractable covariance structure)**.** *For contractable sequences, the covariance* $\mathrm{Cov}(f(X_i), f(X_j))$ *is constant for $i \neq j$.*

Lemma 27 (L^2 bound on window differences)**.** *The L^2 norm of the difference between averages over disjoint windows is bounded.*

Lemma 28 (L^2 contractability bound)**.** *For contractable sequences, certain L^2 norms of block averages are bounded.*

4.2.2 Cesaro Convergence

Lemma 29 (Kallenberg L^2 bound)**.** *The key L^2 bound that drives the Cesaro convergence.*

Lemma 30 (Cesaro to conditional expectation (L^2))**.** *Block averages converge in L^2 to the conditional expectation given the tail.*

Lemma 31 (Cesaro to conditional expectation (L^1))**.** *Block averages converge in L^1 to the conditional expectation given the tail.*

4.2.3 Directing Measure Construction

Lemma 32 (CDF from alpha bounds)**.** *The limiting function* $\alpha(t, \omega)$ *is a.e. in* $[0, 1]$*.*

Lemma 33 (CDF monotonicity)**.** *The limiting function* $\alpha(t, \omega)$ *is monotone in* t *for a.e.* ω*.*

Lemma 34 (Directing measure is probability (L^2 version))**.** *The Stieltjes measure constructed from the limiting CDF is a probability measure a.e.*

Theorem 35 (Weighted sums converge in L^1)**.** *The weighted sums of indicators converge in* L^1*.*

Theorem 36 (Contractable implies Conditionally i.i.d. (via L^2))**.** *If* (X_n) *is contractable, then it is conditionally i.i.d.*

4.3 Via Koopman (Mean Ergodic Theorem)

The Koopman approach uses the Mean Ergodic Theorem via the shift operator on L^2. This is Kallenberg's "first proof" and uses disjoint-block averaging.

4.3.1 Block Averages and Ergodic Theory

Definition 37 (Block average)**.** The *block average* $A_{m,n,k}(f)(\omega) = \frac{1}{n}\sum_{j=0}^{n-1} f(\omega_{k\cdot n+j})$ averages f over the k-th block of size n (indices $[kn, kn + n)$). For $n = 0$, the block average is defined as 0.

Lemma 38 (Koopman-condexp commutation)**.** *The conditional expectation operator onto the shift-invariant subspace commutes with the Koopman operator.*

Theorem 39 (Birkhoff averages converge to condexp)**.** *Birkhoff averages converge in* L^2 *to the conditional expectation given the shift-invariant* σ*-algebra.*

Lemma 40 (Block averages converge in L^1)**.** *For a shift-invariant measure, block averages converge in* L^1 *to the conditional expectation given the shift-invariant* σ*-algebra:* $\int |A_{m,n,k}(f) - \mathbb{E}[f \circ \pi_0 \mid \mathcal{I}]| \, d\mu \to 0$ *as* $n \to \infty$*.*

4.3.2 Contractability and Factorization

Lemma 41 (Conditional expectation lag constancy from exchangeability)**.** *For exchangeable sequences, the conditional expectation of a product does not depend on the lag between coordinates.*

Lemma 42 (Integral product equals block average product)**.** *For contractable sequences, integrals of products factor through block averages.*

Lemma 43 (Product block average L^1 convergence)**.** *Products of block averages converge in* L^1*.*

Theorem 44 (Conditional expectation product factorization)**.** *For contractable sequences, the conditional expectation of a product of indicators factors as a product of conditional expectations.*

Lemma 45 (Bridge from contractability)**.** *For contractable sequences, indicator products satisfy the bridge condition.*

Theorem 46 (Contractable implies Conditionally i.i.d. (via Koopman))**.** *If* (X_n) *is contractable, then it is conditionally i.i.d. This proof uses the Mean Ergodic Theorem via the Koopman operator on* L^2*.*

Chapter 5

Common Ending

All three proofs converge to the same final step: extending from indicators to general sets via a monotone class argument.

Lemma 47 (π-system uniqueness)**.** *Measures on product spaces are determined by their finite-dimensional marginals.*

Theorem 48 (Monotone class theorem)**.** *The monotone class theorem allows extending from indicators to measurable functions.*

Theorem 49 (Conditional independence extension)**.** *Conditional independence on indicators extends to the full product σ-algebra.*

Chapter 6

Main Theorem

Theorem 50 (de Finetti–Ryll-Nardzewski equivalence)**.** *For an infinite sequence* $(X_n)_{n\in\mathbb{N}}$ *of random variables taking values in a* standard Borel space α *(with* α *nonempty), the following are equivalent:*

1. (X_n) *is contractable*
2. (X_n) *is exchangeable*
3. (X_n) *is conditionally i.i.d. (i.e., there exists a directing kernel* ν*)*

Remark: *The martingale proof constructs* ν *from the tail* σ*-algebra* $\mathcal{T}$ *via* $\nu(\omega)(B) = \mathbb{E}[\mathbf{1}_{X_0\in B} \mid \mathcal{T}](\omega)$*.*