Perfect Independence and Stochastic Inequalities

J. Hoffmann-Jørgensen

MaPhySto[1] and the Institute of Mathematical Sciences

University of Aarhus, Denmark

1. Introduction In the 19–seventies, the classical inequalities for absolute sums of independent random variables (Lévy's inequality, Ottaviani's inequality, Jensen's inequality, the symmetrization inequalities, the exponential inequality, the subgaussian inequalities etc.) was extended to norms of sums of independent random vectors taking values in a separable Banach space. In the theory of empirical processes and uniform convergence of stochastic processes, non-measurable norms and non-measurable "random elements" occurs naturally and a crucial part of the theory consists in extending the classical inequalities to vector valued non-measurable functions under various independence and regularity conditions. Note that the usual definition of independence do not apply to non-measurable functions. In the literature, this problem has been bypassed by introducing be the concept "independently defined" (see [3], [15] and Def.2.2 below). The notion of "independently defined" imposes an unpleasant restriction on the underlying probability space and – as we shall see below – is unnecessarily restrictive for the validity of the inequalities. In this paper I shall introduce a series of concepts of independence which applies to arbitrary probability spaces and the objective of this paper is to study the exact form of "independence" under which the classical inequalities hold.

Section 2 is devoted to the study of various concepts of independence of non-measurable functions defined on a probability space $(\Omega, \mathcal{F}, P)$. It is well-known that symmetrization is an important tool for proving the inequalities and Section 3 is devoted to a careful study of "symmetric random elements". In Section 4, we shall study of the validity of the classical inequalities for non-measurable functions under various concepts of independence (actually, it turns that some of the inequalities holds far beyond the realm of "independent random elements"). Working with non-measurable functions, requires that we replace probabilities by outer or inner probabilities and expectations with upper or lower expectations. In Appendix A, you will find a brief summary of the calculus of upper and lower expectations and of upper and lower envelopes. It is well-known that the concept of a perfect map is a major tool to handle upper and lower expectations and upper and lower envelopes. In Appendix B, you will find a brief summary of the fundamental rules of calculus of perfect maps.

We let $\mathbf{R}$ denote *the real line*, we let $\mathbf{N} := \{1, 2, \ldots\}$ denote the set of *positive integers* and we let $\bar{\mathbf{R}} = [-\infty, \infty] := \mathbf{R} \cup \{-\infty, \infty\}$ denote *the extended real line*.

[1] Centre for Mathematical Physics and Stochastics, funded by a grant from the Danish National Research Foundation

In Appendix A, you will the conventional extensions of the arithmetic on the real line to the extended real line. Let T be given set. Then 2^T will denote the set of all subsets of T and a *paving* on T is subset of 2^T. If $\mathcal{A}$ is a paving on T, we let $\sigma(\mathcal{A})$ denote the smallest σ-algebra on T containing $\mathcal{A}$ and we let $\alpha(\mathcal{A})$ denote the smallest algebra on T containing $\mathcal{A}$. We let $\mathbf{R}^T$ and $\bar{\mathbf{R}}^T$ denote the sets of all functions from T into the real line $\mathbf{R}$ and the extended real line $\bar{\mathbf{R}}$, respectively. We let $\mathcal{B}^T := \otimes_{t\in T}\mathcal{B}(\mathbf{R})$ denote the product σ-algebra on $\mathbf{R}^T$. If $f \in \mathbf{R}^T$, we let $||f||_T := \sup_{t\in T}|f(t)|$ denote the *sup-norm* of f and we let $\sup_T f := \sup_{t\in T} f(t)$ and $\inf_T f := \inf_{t\in T} f(t)$ denote *the sup* and *inf-functionals*. Note that the sup-norm $||\cdot||_T$ is $\mathcal{B}^T$-measurable on $\mathbf{R}^T$ if and only if T is countable.

Throughout this paper we let $(\Omega, \mathcal{F}, P)$ denote a fixed probability space and we let $(\Omega, \mathcal{F}_P, \bar{P})$ denote the completion of $(\Omega, \mathcal{F}, P)$; that is, $\mathcal{F}_P$ is the σ-algebra of all P-measurable subsets of Ω and $\bar{P}$ is the probability measure on $(\Omega, \mathcal{F}_P)$ given by $\bar{P}(F) := P^*(F) = P_*(F)$ for all $F \in \mathcal{F}_P$ where $P_*(A)$ and $P^*(A)$ denote the inner and outer probabilities of $A \subseteq \Omega$. If X is a function from Ω into the extended real line $\bar{\mathbf{R}}$, we let E^*X and E_*X denote the upper and lower expectation of X and we let X^* and X_* denote the (a.s. unique) upper and lower P-envelopes of X. If $A \subseteq \Omega$, we let A_* and A^* denote the (a.s. unique) P-kernel and P-hull of A (see Appendix A).

Let $(M, \mathcal{B})$ be a measurable space. Then a measurable function $X : (\Omega, \mathcal{F}_P) \to (M, \mathcal{B})$ will be called a *random function*. If $X : \Omega \to (M, \mathcal{B})$ is a random function, we let P_X denote its *distribution*; that is, the probability measure P_X on $(M, \mathcal{B})$ given by $P_X(B) := \bar{P}(X \in B)$ for all $B \in \mathcal{B}$, and we let $\mathcal{B}_X$ denote the σ-algebra of all P_X-measurable subsets of M. Recall that a random function $X : \Omega \to (M, \mathcal{B})$ is *discrete with discrete range* Q if and only if $Q \in \mathcal{B}$ is countable set such that $P(X \in Q) = 1$ and $\{q\} \in \mathcal{B}$ for all $q \in Q$.

Let $(M, \mathcal{B})$ be a measurable space and let $X, Y : \Omega \to (M, \mathcal{B})$ be random functions with distributions P_X and P_Y. If $P_X(B) = P_Y(B)$ for all $B \in \mathcal{B}$, we say that X *and* Y *have the same distribution on* $(M, \mathcal{B})$. Let $X, Y : \Omega \to M$ be arbitrary functions, then we say that X *and* Y *have the same outer distribution* if $P^*(X \in A) = P^*(Y \in A)$ for all $A \subseteq M$.

2. Independence notions

Let $(M_i, \mathcal{B}_i)$ be a measurable space for all i in the non-empty set I. Then we say that the functions $(X_i : \Omega \to (M_i, \mathcal{B}_i) \mid i \in I)$ are *independent*, if $X_i : \Omega \to (M_i, \mathcal{B}_i)$ is a random function for all $i \in I$ and (X_i) are independent in the usual sense; that is, for every finite non-empty set $\alpha \subseteq I$ we have

$$(2.1) \qquad \bar{P}(X_i \in B_i \ \ \forall\, i \in \alpha) = \prod_{i\in\alpha} \bar{P}(X_i \in B_i) \ \ \forall\,(B_i) \ \text{ satisfying } \ B_i \in \mathcal{B}_i \ \ \forall\, i \in \alpha$$

Note that independence depends heavily on the σ-algebras $\mathcal{B}_i$; for instance, if $\mathcal{T}_i := \{\emptyset, M_i\}$ denotes the trivial σ-algebra for all $i \in I$, then every family of functions $(X_i : \Omega \to (M_i, \mathcal{T}_i) \mid i \in I)$ are independent. If $X_i : \Omega \to M_i$ is an arbitrary function for $i \in I$, then the usual definition of independence do not apply. However, observe that any function $X : \Omega \to M$ is P-measurable with respect to some σ-algebra on M, more precisely:

Let M be a given set and let $X : \Omega \to M$ be a given function without any specified measurability properties. Then we let $\mathcal{M}_P(X)$ denote the largest σ-algebra on M making X P-measurable, and we let $\mathcal{S}_P(X)$ denote the σ-algebra of all P-a.s. X-saturated subsets of Ω; that is,

(2.2) $\mathcal{M}_P(X) := \{ B \subseteq M \mid X^{-1}(B) \in \mathcal{F}_P \}$

(2.3) $\mathcal{S}_P(X) := \{ F \subseteq \Omega \mid \exists\, A \subseteq M \text{ so that } F = X^{-1}(A) \text{ a.s.}\}$

Note that $X : \Omega \to (M, \mathcal{M}_P(X))$ is a random function and that $X : (\Omega, \mathcal{S}_P(X)) \to (M, 2^M)$ is measurable. We let $\hat{P}_X(B) := \bar{P}(X \in B)$ denote the so-called *maximal distribution* of X for all $B \in \mathcal{M}_P(X)$. In our discussion of "independence" of arbitrary functions on Ω, we shall need the following σ-algebras and algebra on Ω:

(2.4) $\mathcal{L}_P(X) := \sigma(\mathcal{K})$, $\mathcal{C}_P(X) := \sigma(\mathcal{S}_P(X) \cup \mathcal{K})$, $\mathcal{A}_P(X) := \alpha(\mathcal{S}_P(X) \cup \mathcal{K})$

where $\mathcal{K}$ denote the set of all P-kernels of all sets $H \in \mathcal{S}_P(X)$

Below you find a series of possible definitions of "independence" of an arbitrary family of functions $(X_i : \Omega \to M_i \mid i \in I)$. Suppose that we have defined a notion of "independence" for families functions $(X_i : \Omega \to M_i \mid\in I)$. Then we say that the notion of "independence" is a *weakly consistent notion of independence* if it satisfying the following two consistency conditions:

(2.A) If the functions $(X_i : \Omega \to M_i \mid i \in I)$ are "independent" and $J \subseteq I$ is a non-empty set, then the functions $(Xi : \Omega \to M_i \mid i \in J)$ are "independent"

(2.B) If the functions $(X_i : \Omega \to M_i \mid i \in I)$ are "independent" and $f_i : M_i \to L_i$ are given functions for all $i \in I$, then the functions $(f_i(X_i) : \Omega \to L_i \mid i \in J)$ are "independent"

and we say that "independence" is a *strongly consistent notion of independence* if "independence" is a weakly consistent notion of independence satisfying the following additional consistency condition:

(2.C) If the functions $(X_i : \Omega \to M_i \mid i \in I)$ are "independent" and $(I_\theta \mid \theta \in \Theta)$ is a non-empty family of non-empty, mutually disjoint subsets of I, then the functions $(X^\theta : \Omega \to M^\theta \mid \theta \in \Theta)$ are "independent" where $M^\theta := \prod_{i \in I_\theta} M_i$ and $X^\theta(\omega) := (X_i(\omega))_{i \in I_\theta}$ for all $\omega \in \Omega$ and all $\theta \in \Theta$

If $G, H \subseteq \Omega$ are given sets, we say that G and H are "independent" if the indicator functions $1_G : \Omega \to \{0,1\}$ and $1_H : \Omega \to \{0,1\}$ are "independent" and if $\mathcal{G}$ and $\mathcal{H}$ are pavings on Ω, we say $\mathcal{G}$ and $\mathcal{H}$ are "independent" if G and H are "independent" for all $G \in \mathcal{G}$ and all $H \in \mathcal{H}$.

Lemma 2.1: *Let L and M be a given sets and let $X : \Omega \to M$ and $Y : \Omega \to L$ be a given functions such that $Y = h(X)$ a.s. for some function $h : M \to L$. If $\mathcal{K}$ denotes the paving of all P-kernels of all sets $H \in \mathcal{S}_P(X)$, then we have*

(1) $\mathcal{S}_P(Y) \subseteq \mathcal{S}_P(X)$, $X^{-1}(\mathcal{M}_P(X)) \subseteq \mathcal{K} \subseteq \mathcal{L}_P(X) \subseteq \mathcal{F}_P$, $X(\Omega) \in \mathcal{M}_P(X)$

(2) $\mathcal{L}_P(Y) \subseteq \mathcal{L}_P(X)$, $\mathcal{L}_P(Y) \subseteq \mathcal{L}_P(X)$, $\mathcal{C}_P(Y) \subseteq \mathcal{C}_P(X)$, $\mathcal{A}_P(Y) \subseteq \mathcal{A}_P(X)$

(3) $A \in \mathcal{M}_P(Y) \quad \Leftrightarrow \quad h^{-1}(A) \in \mathcal{M}_P(X)$

(4) *$\mathcal{K}$ is stable under countable intersections and we have $\mathcal{F}_P \cap \mathcal{S}_P(X) \subseteq \mathcal{K}$*

(5) *If $f : M \to \bar{\mathbf{R}}$ is a given function and $f(X)^*$ and $f(X)_*$ are any given upper and lower P-envelopes of $f(X)$, then $f(X)^*$ and $f(X)_*$ are $\mathcal{L}_P(X)$-measurable and we have*

  (a) $\{ f(X)^* \leq s\} \in \mathcal{K}$ *and* $\{ f(X)_* \geq s\} \in \mathcal{K} \quad \forall s \in \mathbf{R}$

(6) *$(M, \mathcal{M}_P(X), \hat{P}_X)$ is a complete probability space and if $\mathcal{B}$ is a σ-algebra on M such that $X : \Omega \to (M, \mathcal{B})$ is a random function with distribution P_X , then we have*

  (a) $\mathcal{B}_X \subseteq \mathcal{M}_P(X)$ *and* $\bar{P}_X(B) = \hat{P}_X(B) \;\; \forall B \in \mathcal{B}_X$

(7) *If $\mathcal{B}$ is a σ-algebra on M such that $X : \Omega \to (M, \mathcal{B})$ is a $(P, \mathcal{B})$-perfect function, then we have*

  (a) $\mathcal{L}_P(X) = \mathcal{F}_P \cap \mathcal{S}_P(X)$, $\mathcal{S}_P(X) = \mathcal{C}_P(X)$ *and* $\mathcal{M}_P(X) = \mathcal{B}_X$

Proof: By Lem.A.5.9, we see that $\mathcal{K}$ is stable under countable intersections and since $Y^{-1}(A) = X^{-1}(h^{-1}(A))$ a.s. for all $A \subseteq L$ and $F = F_*$ for all $F \in \mathcal{F}_P$, we see that (1)–(4) hold. By Lem.A.5.5+6, we see that (5.a) holds and since $\mathcal{L}_P(X) = \sigma(\mathcal{K})$, we see that $f(X)^*$ and $f(X)_*$ are $\mathcal{L}_P(X)$-measurable. (6) is an immediate consequence of the definition of the maximal distribution, and (7) follows easily from Lem.B.1. $\square$

Definition 2.2: (Complete independence) Let $(M_i \mid i \in I)$ be a non-empty family of sets and let $(X_i : \Omega \to M_i \mid i \in I)$ be a given family of functions. Then $X_i : \Omega \to (M_i, \mathcal{M}_P(X_i))$ is a random function for all $i \in I$ and we say that the functions $(X_i : \Omega \to M_i \mid i \in I)$ are *completely independent* if the random functions $(X_i : \Omega \to (M_i, \mathcal{M}_P(X_i)) \mid i \in I)$ are (ordinary) independent. Since $\mathcal{B}_i \subseteq \mathcal{M}_P(X_i)$ for every σ-algebra on M_i making X_i P-measurable, we see that complete independence implies ordinary independence (but the converse implication fails in general). By Lem.2.1.3, we see that the notion "completely independent" is weakly consistent but the following example shows that "completely independent" is *not* strongly consistent.

Example: λ denote the Lebesgue measure on the Borel σ-algebra $\mathcal{B}$ of unit interval $I = [0, 1]$. Let $A \subseteq I$ be a given set satisfying $\lambda^*(A) = 1$ and $\lambda_*(A) = 0$.

If $A^c := I \setminus A$ denotes the complement of A, we have $\lambda^*(A^c) = 1$ and $\lambda_*(A^c)$. Let $a \in A$ and $b \in A^c$ be given points and let us define

$$X_1(\omega) = \omega\, 1_A(\omega) + a\, 1_{A^c}(\omega) \ \ , \ \ X_2(\omega) = \omega\, 1_{A^c}(\omega) + b\, 1_A(\omega) \ \ , \ \ X_3(\omega) = \omega$$

for all $\omega \in I$. Then X_1, X_2, X_3 are functions from I into I. Let $B \subseteq I$ be given. Since $X_1^{-1}(B) = A \cap B$ if $a \notin B$ and $X_1^{-1}(B) = (A \cap B) \cup A^c$ if $a \in B$, we see that either $X_1^{-1}(B) \subseteq A$ or $X_1^{-1}(B) \supseteq A^c$ and since $\lambda_*(A) = 0$ and $\lambda^*(A^c) = 1$, we see that $\bar{\lambda}(F) = 0$ or 1 for all $F \in X_1^{-1}(\mathcal{M}_\lambda(X_1)$. In the same manner we see that $\bar{\lambda}(F) = 0$ or 1 for all $F \in X_2^{-1}(\mathcal{M}_\lambda(X_2)$. Hence, we see that X_1, X_2, X_3 are completely independent. Let us define $U := (X_1, X_2)$. Then Since U is injective on $I \setminus \{a, b\}$, we have $U^{-1}(\mathcal{M}_\lambda(U)) = \mathcal{B}_\lambda$ and since $X_3^{-1}(\mathcal{M}_\lambda(X_3)) = \mathcal{B}_\lambda$, we see that (X_1, X_2) and X_3 are *not* completely independent.

Definition 2.3: (Independently defined functions) Let M_i be a given of set for all i in the non-empty set I. Then we say that the functions $(X_i : \Omega \to M_i \mid i \in I)$ are *independently defined* if and only if the exists probability spaces $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$ and $(\Omega_i, \mathcal{F}_i, P_i)$ and functions $\hat{X}_i : \Omega_i \to M_i$ for $i \in I$ satisfying

(1) $\quad \Omega = \tilde{\Omega} \times \prod_{i \in I} \Omega_i \ \ , \ \ \mathcal{F} = \tilde{\mathcal{F}} \otimes \bigotimes_{i \in I} \mathcal{F}_i \ \ , \ \ P = \tilde{P} \otimes \bigotimes_{i \in I} P_i$

(2) $\quad X_i(\omega) = \hat{X}_i(\omega_i) \quad \forall\, \omega = (\tilde{\omega}, (\omega_\nu)_{\nu \in I}) \in \Omega \ \forall\, i \in I$

Observe that the concept "independently defined" imposes an unpleasant restriction on the form of the underlying probability space $(\Omega, \mathcal{F}, P)$ and the in next two definitions we shall we shall remove this restriction. By the definition of "independently defined" it is obvious that "independently defined" is strongly consistent.

Definition 2.4: (Perfect independence) Let M_i be a given set for all i in the non-empty set I. Then we say that the functions $(X_i : \Omega \to M_i \mid i \in I)$ are *perfectly independent* if and only if the exist measurable spaces $(\Omega_i, \mathcal{F}_i)$ and functions $\xi_i : \Omega \to \Omega_i$ and $\hat{X}_i : \Omega_i \to M_i$ satisfying (see Appendix B):

(1) The functions $(\xi_i : \Omega \to (\Omega_i, \mathcal{F}_i) \mid i \in I)$ are independent and we have $X_i(\omega) = \hat{X}_i(\xi_i(\omega))$ for all $\omega \in \Omega$ and all $i \in I$

(2) The function $\xi(\omega) := (\xi_i(\omega))_{i \in I}$ from Ω into the product space $(\Omega^I, \mathcal{F}^I)$ is $(P, \mathcal{F}^I)$-perfect where $\Omega^I := \prod_{i \in I} \Omega_i$ and $\mathcal{F}^I := \otimes_{i \in I} \mathcal{F}_i$

If (1) holds and the function $\xi(\omega)$ in (2) is strictly $(P, \mathcal{F}^I)$-perfect, we say that the functions (X_i) are *strictly perfectly independent*. Below, we shall see that "perfectly independent" and "strictly perfectly independent" are strongly consistent.

Definition 2.5: (Total independence) Let L and M be two given sets and let $X:\Omega\to L$ and $Y:\Omega\to M$ be given functions. If $\psi:L\times M\to\bar{\mathbf{R}}$ is a given function and $F\subseteq\Omega$ is a given set, we say that X and Y are *upper ψ-independent* if

(1) $E^*\psi(X_1,X_2)\}\geq E^*\psi^\circ(X_1)$ where $\psi^\circ(x):=E^*\psi(x,X_2)\ \forall\, x\in M$

If X and Y are upper 1_D-independent for all $D\subseteq L\times M$, we say that X and Y are *totally independent*. If $(L,\mathcal{A})$ and $(M,\mathcal{B})$ are measurable spaces, then it follows easily that we have

(2) If $X:\Omega\to(L,\mathcal{A})$ and $Y:\Omega\to(M,\mathcal{B})$ are independent random functions, then X and Y are upper ψ-independent for every $P_{X,Y}$-measurable function $\psi : L\times M \to \bar{\mathbf{R}}$

Note that the notion "totally independent" is non-symmetric in (X,Y) and I do not know if the implication "X and Y are totally independent" $\Rightarrow$ "Y and X are totally independent" holds in general. However, if M and L are countable sets, then Thm.2.8 shows that this indeed so. If X and Y are totally independent and Y and X are totally independent, we say that X,Y are totally independent.

Let M_i be a given set and let $X_i:\Omega\to M_i$ be a given function for all i in the non-empty set I. If $\alpha\subseteq I$ is non-empty subset of I, we let $M^\alpha:=\prod_{i\in\alpha}M_i$ denote the product space and we let $X^\alpha(\omega):=(X_i(\omega))_{i\in\alpha}$ denote *the α-marginal* of X. Then X^α is a function from Ω into the product space M^α and we say that the functions $(X_i:\Omega\to M_i\mid i\in I)$ are *totally independent* if X^α and X^β are totally independent whenever α and β are disjoint non-empty subsets of I. Below, we shall see that "totally independent" is strongly consistent.

Theorem 2.6: *Let L_1, M_1, L_2 and M_2 be given sets and let $X_1:\Omega\to M_1$, $X_2:\Omega\to M_2$, $Y_1:\Omega\to L_1$ and $Y_2:\Omega\to L_2$ be given functions such that $Y_i=f_i(X_i)$ a.s. where $f_i:M_i\to L_i$ are given functions for $i=1,2$. Let $(S_1,\mathcal{A}_1)$ and $(S_2,\mathcal{A}_2)$ be countably separated measurable spaces and let $Z_i:(\Omega,\mathcal{C}_P(X_i))\to(S_i,\mathcal{A}_i)$ be measurable functions for $i=1,2$. Then we have*

(1) *If X_1 and X_2 are totally independent, then Y_1 and Y_2 are totally independent and upper ϕ-independent for every function $\phi:L_1\times L_2\to\bar{\mathbf{R}}$*

(2) *If X_1 and X_2 are perfectly independent, then (Z_1,Y_1) and (Z_2,Y_2) are totally independent*

(3) *If $(1_F,X_1)$ and X_2 are totally independent and $\psi:M_1\times M_2\to\bar{\mathbf{R}}$ is a given function, then we have*

   (a) $E^*\{1_F\,\psi(X_1,X_2)\}\geq E^*\{1_F\,\psi^\circ(X_1)\}$ *where* $\psi^\circ(x):=E^*\psi(x,X_2)$

Let I be a non-empty set, let M_i be a given set for all $i \in I$ and let the functions $(X_i : \Omega \to M_i \mid i \in I)$ be perfectly independent. If $\alpha \subseteq I$ is a non-empty countable set and $A_i \subseteq M_i$ are any given sets for all $i \in \alpha$, then we have

(4) $\quad P^*(X_i \in A_i \quad \forall\, i \in \alpha) = \prod_{i\in\alpha} P^*(X_i \in A_i)$

(5) $\quad P_*(X_i \in A_i \quad \forall\, i \in \alpha) = \prod_{i\in\alpha} P_*(X_i \in A_i)$

Proof: (1): Let $\phi : L_1 \times L_2 \to [0,\infty]$ be a non-negative given function and let us define $\psi(x,y) := \phi(f_1(x), f_2(y))$, $\phi^\circ(u) := E^*\phi(u, Y_2)$ and $\psi^\circ(x) := E^*\psi(x, X_2)$ for all $(x,y) \in M_1 \times M_2$ and all $u \in L_1$. Since $Y_i = f_i(X_i)$ a.s., we have $\phi(Y_1,Y_2) = \psi(X_1,X_2)$ a.s. and $\phi^\circ(Y_1) = \psi^\circ(X_1)$ a.s. and if we define $D_t := \{(x,y) \in M_1 \times M_2 \mid \psi(x,y) > t\}$ and $f(x,t) := P^*((x,Y) \in D_t)$ for all $x \in M_1$ and all $t \geq 0$, then by Thm.A.3.3 we have

$$E^*\phi(Y_1,Y_2) = E^*\psi(X_1,X_2) = \int_0^\infty P^*(\,(X_1,X_2) \in D_t)\,dt$$

$$\psi^\circ(x) = E^*\psi(x,X_1) = \int_0^\infty f(x,t)\,dt \quad \forall\, x \in M$$

Since X_1 and X_2 are totally independent, we have $P^*((X_1,X_2) \in D_t) \geq E^* f(X_1,t)$ and since $f(X(\omega),\cdot)$ is decreasing and non-negative for all $\omega \in \Omega$, then by Thm.A.7 we have

$$\begin{aligned} E^*\phi(Y_1,Y_2) &= \int_0^\infty P^*(\,(X_1,X_2) \in D_t)\,dt \geq \int_0^\infty dt \int^* f(X_1(\omega),t)\,P(d\omega) \\ &\geq \int^* P(d\omega) \int_0^\infty f(X_1(\omega),t)\,dt = \int^* \psi^\circ(X_1(\omega))\,P(d\omega) \\ &= E^*\psi^\circ(X_1) = E^*\phi^\circ(Y_1) \end{aligned}$$

Hence, we see that Y_1 and Y_2 are upper ϕ-independent for every non-negative function $\phi : L_1 \times L_2 \to [0,\infty]$. Adding a real constant to ψ do not affect the equation derived above (see Lem.A.2.10). Hence, if $\phi : L_1 \times L_2 \to \bar{\mathbf{R}}$ is an arbitrary function, we see that Y_1 and Y_2 are upper ϕ_a-independent for every $a \in \mathbf{R}$ where $\phi_a := \phi \vee a$. Let us define $\phi^\circ(u) := E^*\phi(u,Y_2)$ and $\phi_a^\circ(u) := E^*\phi_a(u,Y_2)$ for all $u \in L_1$. Then we have $\phi^\circ \leq \phi_a^\circ$ and since Y_1 and Y_2 are upper ϕ_a-independent, we have

$$E^*(\phi(Y_1,Y_2) \vee a) \geq E^*\phi_a^\circ(Y_1) \geq E^*\phi^\circ(Y_1) \quad \forall\, a \in \mathbf{R}$$

Letting $a \downarrow -\infty$, then by Lem.A.2.3 we see that Y_1 and Y_2 are upper ϕ-independent for all $\phi : L_1 \times L_2 \to \bar{\mathbf{R}}$. In particular, we see that Y_1 and Y_2 are totally independent.

(2): Suppose that X_1 and X_2 are perfectly independent and let $(\Omega_i, \mathcal{F}_i, \xi_i, \hat{X}_i)$ be chosen according to Def.2.4.1+2 for $i = 1,2$. By Def.2.4.2, we have that $\xi := (\xi_1, \xi_2)$ is $(P, \mathcal{F}_1 \otimes \mathcal{F}_2)$-perfect and by Def.2.4.1 we have that $P_\xi = P_{\xi_1} \otimes P_{\xi_2}$. So by

Lem.B.6.1 and Lem.B.7.1 we see that ξ_i is $(P, \mathcal{F}_i)$-perfect for $i = 1, 2$ and since $X_1(\omega) = \hat{X}_1(\xi_1(\omega))$, then by Lem.2.1.2+7 we have $\mathcal{C}_P(X_i) \subseteq \mathcal{S}_P(\xi_i)$. Since $(S_i, \mathcal{A}_i)$ is countable separated, there exist sets $A_1^i, A_2,^i, \ldots \in \mathcal{A}_i$ separating points in L_i (i.e. if $u \neq v \in L_i$, then there exists an integer $n \geq 1$ such that $u \in A_n^i$ and $v \notin A_n^i$). Since $Z_i : (\Omega, \mathcal{C}_P(X_i)) \to (S_i, \mathcal{A}_i)$ is measurable and $\mathcal{C}_P(X_i) \subseteq \mathcal{S}_P(\xi_i)$, there exist a P-null set N and sets $F_n^i \subseteq \Omega_i$ such that $Z_i^{-1}(A_n^i) \setminus N = \xi_i^{-1}(F_n^i) \setminus N$ for all $n \geq 1$. Let $\omega, \omega' \in \Omega \setminus N$ be given points such that $Z_i(\omega) \neq Z_i(\omega')$. Since $A_1^i, A_2^i, \ldots$ separates points in L_i, then exists an integer $n \geq 1$ such that $\omega \in Z_i^{-1}(A_n^i) = \xi_i^{-1}(F_n^i)$ and $\omega' \notin Z_i^{-1}(A_n^i) = \xi_i^{-1}(F_n^i)$. Hence, we see that $\xi_i(\omega) \neq \xi_i(\omega')$ for all $\omega, \omega' \in \Omega \setminus N$ satisfying $Z_i(\omega) \neq Z_i(\omega')$ and consequently, there exists a function $h_i : \Omega_i \to L_i$ such that $Z_i(\omega) = h_i(\xi_i(\omega))$ for all $\omega \in \Omega \setminus N$. So by Def.2.4.1 we have $(Z_i(\omega), Y_i(\omega)) = (h_i(\xi_i(\omega), f_i(\hat{X}_i(\xi_i(\omega))))$ a.s. for $i = 1, 2$. Hence, by (1) it suffices to show that $\xi_1 : \Omega \to \Omega_1$ and $\xi_2 : \Omega \to \Omega_2$ are totally independent. So let $C \subseteq \Omega_1 \times \Omega_2$ be a given set and let us define $\psi(\omega_1) := P^*((\omega_1, \xi_2) \in C)$ for all $\omega_1 \in \Omega_1$. Since ξ and ξ_1 are perfect and $P_\xi = P_{\xi_1} \otimes P_{\xi_2}$, then by Thm.A.6.1 and Lem.B.1.4 we have

$$P^*(\,(\xi_1, \xi_2) \in C) = P_\xi^*(C) \geq \int^* \phi(\omega_1)\, P_{\xi_1}(d\omega_1) = E^* \phi(\xi_1)$$

that is, ξ_1 and ξ_2 are totally independent which completes the proof of (2).

(3): Let us define $\phi((\delta, x), y) := \delta\, \psi(x, y)$ for all $(\delta, x) \in \{0, 1\} \times M_1$ and all $y \in M_2$. By (1), we have that $(1_F, X_1)$ and X_2 are upper ϕ-independent and since $E^* \phi(\delta, x), X_2) = \delta\, \psi^\circ(x)$, we see that (3.a) holds.

(4+5): Let $\alpha \subseteq I$ be a given countable set. Since the functions $(X_i \mid i \in I)$ are perfectly independent, then by Lem.B.6.1 and Lem.B.7.1 we see that the functions $(X_i \mid i \in \alpha)$ are perfectly independent. Hence, it suffices to prove (4+5) when $\alpha = I$ and I is countable. So suppose that I is countable and let $A_i \subseteq M_i$ be given sets for all $i \in I$. Let $(\Omega_i, \mathcal{F}_i, \xi_i, \hat{X}_i)$ be chosen according to Def.2.4.1+2 and let us define $(\Omega^I, \mathcal{F}^I) := (\prod_{i \in I} \Omega_i, \otimes_{i \in I} \mathcal{F}_i)$ and $\xi(\omega) := (\xi_i(\omega))_{i \in I}$. Let us define $B_i := \hat{X}_i^{-1}(A_i)$ for all $i \in I$ and let P_ξ denote the distribution of ξ. By Def.2.4.1, we have $P_\xi = \otimes_{i \in I} P_{\xi_i}$ and by Def.2.4.1 we have $\{X_i \in A_i \ \forall\, i \in I\} = \xi^{-1}(\prod_{i \in I} B_i)$ and $\{X_i \in A_i\} = \xi_i^{-1}(B_i)$ for all $i \in I$. Since ξ is $(P, \mathcal{F}^I)$-perfect, then by Lem.B.6.1 and Lem.B.7.1 we see that ξ_i is $(P, \mathcal{F}_i)$-perfect for all $i \in I$ and so by Thm.A.8 we have

$$P^*(X_i \in A_i \ \ \forall\, i \in I) = P_\xi^*\left(\prod_{i \in I} B_i\right) = \prod_{i \in I} P_{\xi_i}^*(B_i) = \prod_{i \in I} P^*(X_i \in A_i)$$

$$P_*(X_i \in A_i \ \ \forall\, i \in I) = (P_\xi)_*\left(\prod_{i \in I} B_i\right) = \prod_{i \in I} (P_{\xi_i})_*(B_i) = \prod_{i \in I} P_*(X_i \in A_i)$$

which proves (4) and (5). $\square$

Lemma 2.7: *Let* $G \subseteq \Omega$ *and* $H \subseteq \Omega$ *be totally independent sets and let us define* $G_0 := \Omega \backslash G$ *,* $G_1 := A$ *,* $H_0 := \Omega \backslash H$ *and* $H_1 := H$ *. Then for all* $\epsilon, \delta \in \{0, 1\}$ *we have*

(1) $\quad P(G_\epsilon^*)\, P(H_\delta^*) = P^*(G_\epsilon)\, P^*(H_\delta) = P^*(G_\epsilon \cap H_\delta) = P(G_\epsilon^* \cap H_\delta^*)$

(2) $\quad P(G_{\epsilon *})\, P(H_{\delta *}) = P_*(G_\epsilon)\, P_*(H_\delta) = P_*(G_\epsilon \cap H_\delta) = P(G_{\epsilon *} \cap H_{\delta *})$

(3) $\quad (G_\epsilon \cap H_\delta)^* = G_\epsilon^* \cap H_\delta^*$ *a.s. and* $(G_\epsilon \cup H_\delta)_* = G_{\epsilon *} \cup H_{\delta *}$ *a.s.*

where A_* *and* A^* *denote the (a.s. unique)* P*-kernel and* P*-hull of* $A \subseteq \Omega$ *.*

Proof: (1): By Thm.2.6.1, we have that 1_{G_ϵ} and 1_{H_δ} are totally independent for all $\epsilon, \delta \in \{0, 1\}$. Hence, it suffices to prove (1) when $\epsilon = \delta = 1$. The first equality in (1) is evident. Let us define $D := \{(1, 1)\}$. Then we have $P^*(G \cap H) = P^*((1_G, 1_H) \in D)$ and $P^*((\delta, 1_H) \in D) = \delta\, P^*(H)$ for $\delta = 0, 1$ and since G and H are totally independent, we have

(*) $\qquad P^*(G \cap H) = P^*(\,(1_G, 1_H) \in D) \geq E^*(1_G\, P^*(H)) = P^*(G)\, P^*(H)$

Let us define $D := \{(1, 0), (1, 1), (0, 1)\}$ and $\phi(\delta) := P^*((\delta, 1_H) \in D)$ for all $\delta = 0, 1$. Then we have $G \cup H = \{(1_G, 1_H) \in D\}$, $\phi(0) = P^*(H)$ and $\phi(1) = 1$. Hence, we have

$$\phi(1_G) = P^*(H)\, 1_{G^c} + 1_G = P^*(H) + (1 - P^*(H))\, 1_G$$

and so and Lem.A.2.5+10 and total independence we have

$$\begin{aligned} P^*(G \cup H) &= P^*(\,(1_G, 1_H) \in D) \geq E^* \phi(1_G) = P^*(H) + (1 - P^*(H))\, P^*(G) \\ &= P^*(H) + P^*(G) - P^*(G)\, P^*(H) \end{aligned}$$

Hence, by (*) and Lem.A.1.5 we have

$$\begin{aligned} P^*(G \cup H) + P^*(G \cap H) &\geq P^*(G \cup H) + P^*(G)\, P^*(H) \\ &\geq P^*(G) + P^*(H) \geq P^*(G \cup H) + P^*(G \cap H) \end{aligned}$$

Hence, we see that $P^*(G \cup H) + P^*(G \cap H) = P^*(G) + P^*(H)$ and that $P^*(G \cap H) = P^*(G)\, P^*(H)$. In particular, we see that the second inequality in (2) holds and by Lem.A.5.8 we have $(G \cup H)^* = G^* \cup H^*$ a.s. Hence, we have

$$\begin{aligned} P^*(G) + P^*(H) &= P^*(G \cup H) + P^*(G \cap H) = P(G^*) + P(H^*) \\ &= P(G^* \cup H^*) + P(G^* \cap H^*) = P^*(G \cup H) + P(G^* \cap H^*) \end{aligned}$$

Thus, we see that $P^*(G \cap H) = P(G^* \cap H^*)$ which completes the proof of (1).

(2): As in the proof of (1), we see that it suffices to prove (2) when $\epsilon = \delta = 1$. The first equality in (2) is evident and the last equality follows from Lem.A.5.9. In the proof of (1) we found that $P^*(G_1 \cup H_1) + P^*(G_1 \cap H_1) = P^*(G_1) + P^*(H_1)$. So by (1) and Lem.A.1.1 we have

$$
\begin{aligned}
P_*(G \cap H) &= 1 - P^*(G_1 \cup H_1) = 1 - \{P^*(G_1) + P^*(H_1) - P^*(G_1 \cap H_1)\} \\
&= 1 + P^*(G_1)\, P^*(H_1) - P^*(G_1) - P^*(H_1) \\
&= P_*(G) - P^*(H_1)\, P_*(G) = P_*(G)\, P_*(H)
\end{aligned}
$$

which completes the proof of (2).

(3): Since $G_\epsilon \cap H_\delta \subseteq G_\epsilon^* \cap H_\delta^*$ a.s., we see that the first equality in (3) follows from the last equality in (1). The last equality in (3) follows from the first equality and Lem.A.5.6. $\square$

Theroem 2.8: *Let M_1 and M_2 be non-empty sets and let $X_1 : \Omega \to M_1$ and $X_2 : \Omega \to M_2$ be given function. Let C^* denote a fixed P-hull of C whenever $C \subseteq \Omega$. If M_1 and M_2 are finite, then the following five statements are equivalent:*

(1) X_1 *and* X_2 *are totally independent*

(2) $\{X_1 \in A\}$ *and* $\{X_2 \in B\}$ *are totally independent* $\forall\, A \subseteq M_1\ \forall\, B \subseteq M_2$

(3) $\{X_1 \in A\}^*$ *and* $\{X_2 \in B\}^*$ *are independent* $\forall\, A \subseteq M_1\ \forall\, B \subseteq M_2$ *and* $\{X_1 = u\}^* \cap \{X_2 = v\}^* = \{X_1 = u\ ,\ X_2 = v\}^*$ *a.s.* $\forall\, (u,v) \in M_1 \times M_2$

(4) X_1 *and* X_2 *are strictly perfectly independent*

(5) X_1, X_2 *are totally independent*

and if M_1 and M_2 are countable, then the four statements (1), (2), (3) and (5) are equivalent. Moreover, if M_1 and M_2 are arbitrary sets, then we have

(6) X_1 *and* X_2 *are totally independent, then* X_1 *and* X_2 *are completely independent*

(7) *If $X_1 : \Omega \to (M_1, 2^{M_1})$ and $X_2 : \Omega \to (M_2, 2^{M_2})$ are independent P-measurable functions and one of the sets M_1 or M_2 is not real measurable, then X_1 and X_2 are perfectly independent*

Remark: Recall that a set M is *real measurable* if and only if there exists a probability measure μ on $(M, 2^M)$ such that $\mu(\{x\}) = 0$ for all $x \in M$. If M is real measurable, then it follows easily that every set L with $\operatorname{card} L \geq \operatorname{card} M$ is real measurable. Hence if RM denote the smallest cardinality of a real measurable set, then M is real measurable if and only if $\operatorname{card} M \geq \mathrm{RM}$. It is well-known that RM is a very large cardinal – actually so large that it need not exists in the sense that it is consistent with the usual axioms of set theory (including the axiom of choice) to assume that there exists no real measurable sets at all.

Proof: Suppose that M_1 and M_2 are finite and let me prove the equivalence of (1)–(5).

(1) $\Rightarrow$ (2): Follows from Thm.2.6.1.

(2) $\Rightarrow$ (3): Follows from Lem.2.7.2+4.

(3) $\Rightarrow$ (4): Suppose that (3) holds and let $\epsilon \in \{1,2\}$ be given and fixed for a moment. Let $u \in M_\epsilon$ be given and let D_u^ϵ be a given P-hull of $X_\epsilon^{-1}(u)$ satisfying $D_u^\epsilon \in \mathcal{F}$ and $D_u^\epsilon \supseteq X_\epsilon^{-1}(u)$. Let $\mathcal{F}_\epsilon$ denote the σ-algebra generated by the sets $(D_u^\epsilon \mid u \in M_\epsilon)$ and let $(F_i^\epsilon \mid i \in L_\epsilon)$ denote the non-empty atoms of the σ-algebra $\mathcal{F}_\epsilon$. Since M_ϵ is finite, then $\mathcal{F}_\epsilon$ is finite and so we see that L_ϵ is a non-empty finite set and that $F_i^\epsilon \in \mathcal{F}_\epsilon \subseteq \mathcal{F}$ for all $i \in L_\epsilon$. Let $u \in M_\epsilon$ be given and let us define $\pi_u^\epsilon := \{i \in L_\epsilon \mid F_i^\epsilon \cap D_u^\epsilon \neq \emptyset\}$. Since $D_u^\epsilon \in \mathcal{F}_\epsilon$, we have $D_u^\epsilon = \cup_{i\in\pi_u^\epsilon} F_i^\epsilon$ with the usual convention that an empty union is empty. Let us define

$$\Gamma_\epsilon := \{\,(i,u) \in L_\epsilon \times M_\epsilon \mid i \in \pi_u^\epsilon\} \;\;,\;\; \mathcal{G}_\epsilon := \{\,\Gamma_\epsilon \cap (\,\pi \times M_\epsilon) \mid \pi \subseteq L_\epsilon\}$$

$$\hat{X}_\epsilon(i,u) := u \;\;,\;\; A_\gamma^\epsilon := F_i^\epsilon \cap X_\epsilon^{-1}(u) \;\;,\;\; \Lambda_j^\epsilon := \Gamma_\epsilon \cap (\{j\} \times M_\epsilon)$$

for all $\gamma = (i,u) \in \Gamma_\epsilon$ and all $j \in L_\epsilon$. Then $(\Gamma_\epsilon, \mathcal{G}_\epsilon)$ is a σ-discrete measurable space and since $D_u^\epsilon \supseteq X_\epsilon^{-1}(u)$ for all $u \in M_\epsilon$, we have $\Omega = \cup_{u\in M_\epsilon} D_u^\epsilon$. Let $i \in L_\epsilon$ be given. Since $F_i^\epsilon \neq \emptyset$, there exists $u \in M_\epsilon$ such that $F_i^\epsilon \cap D_u^\epsilon \neq \emptyset$ or equivalently $i \in \pi_u^\epsilon$. Hence, we see that $(i,u) \in \Lambda_i^\epsilon$; that is, $\Lambda_i^\epsilon \neq \emptyset$ for all $i \in L_\epsilon$. So by the definition of the σ-algebra $\mathcal{G}_\epsilon$, we see that $(\Lambda_i^\epsilon \mid i \in L_\epsilon)$ is the set of all non-empty $\mathcal{G}_\epsilon$-atoms. Since $D_u^\epsilon = \cup_{i\in\pi_u} F_i^\epsilon \supseteq X_\epsilon^{-1}(u)$ for all $u \in M_\epsilon$, we have

$$\bigcup_{\gamma\in\Gamma_\epsilon} A_\gamma^\epsilon = \bigcup_{u\in M_\epsilon} \bigcup_{i\in\pi_u} F_i^\epsilon \cap X_\epsilon^{-1}(u) = \bigcup_{u\in M_\epsilon} \left\{ X_\epsilon^{-1}(u) \cap \bigcup_{i\in\pi_u} F_i^\epsilon \right\}$$

$$= \bigcup_{u\in M_\epsilon} X_\epsilon^{-1}(u) \cap D_u^\epsilon = \bigcup_{u\in M_\epsilon} X_\epsilon^{-1}(u) = \Omega$$

and since the sets $(F_i^\epsilon \mid i \in L_\epsilon)$ are mutually disjoint and the sets $(X_\epsilon^{-1}(u) \mid u \in M_\epsilon)$ are mutually disjoint, we see that $(A_\gamma^\epsilon \mid \gamma \in \Gamma_\epsilon)$ is a disjoint partition of Ω . Hence, we may define $\xi_\epsilon : \Omega \to \Gamma_\epsilon$ as follows $\xi_\epsilon(\omega) := \gamma$ if γ is the unique $\gamma \in \Gamma_\epsilon$ satisfying $\omega \in A_\gamma^\epsilon$. Since $X_\epsilon(\omega) = u$ for all $\omega \in A_{(i,u)}^\epsilon$, we see that $X_\epsilon(\omega) = \hat{X}_\epsilon(\xi_\epsilon(\omega))$ for all $\omega \in \Omega$.

By the definition of ξ_ϵ , we have $\xi_\epsilon^{-1}(\gamma) = A_\gamma^\epsilon$ for all $\gamma \in \Gamma_\epsilon$. Let $i \in L_\epsilon$ be given and let us define $C_i^\epsilon := \{u \in M_\epsilon \mid i \in \pi_u^\epsilon\}$. Let $u \in X_\epsilon(F_i^\epsilon)$ be given. Then we have $F_i^\epsilon \cap X_\epsilon^{-1}(u) \neq \emptyset$ and since $X_\epsilon^{-1}(u) \subseteq D_u^\epsilon$, we see that $F_i^\epsilon \cap D_u^\epsilon \neq \emptyset$; that is, $i \in \pi_u^\epsilon$ or equivalently $u \in C_i^\epsilon$. So we see that $X_\epsilon(F_i^\epsilon) \subseteq C_i^\epsilon$ and consequently, we have $F_i^\epsilon \subseteq X_\epsilon^{-1}(X_\epsilon(F_i^\epsilon)) \subseteq X_\epsilon^{-1}(C_i)$. Hence, by the definition of Λ_i^ϵ , we have

$$\xi_\epsilon^{-1}(\Lambda_i^\epsilon) = \bigcup_{u\in C_i^\epsilon} F_i^\epsilon \cap X_\epsilon^{-1}(u) = F_i^\epsilon \cap X^{-1}(C_i^\epsilon) = F_i^\epsilon \in \mathcal{F}_\epsilon \;\; \forall\, i \in L_\epsilon$$

Since $(\Lambda_i^\epsilon \mid i \in L_\epsilon)$ are the atoms of $\mathcal{G}_\epsilon$, we have

(i) $\quad \xi_\epsilon : (\Omega, \mathcal{F}_\epsilon) \to (\Gamma_\epsilon, \mathcal{G}_\epsilon)$ is a measurable function

(ii) $\xi_\epsilon^{-1}(\gamma) = A_\gamma^\epsilon = F_i^\epsilon \cap X_\epsilon^{-1}(u) \ \ \forall\, \gamma = (i,u) \in \Gamma_\epsilon$ and $\xi_\epsilon^{-1}(\Lambda_i^\epsilon) = F_i^\epsilon \ \ \forall\, i \in L_\epsilon$

(iii) $X_\epsilon(\omega) = \hat{X}_\epsilon(\xi_\epsilon(\omega)) \quad \forall\, \omega \in \Omega$

Let $\gamma = (i,u) \in \Gamma_\epsilon$ be given. Since $i \in \pi_u^\epsilon$, we have $F_i^u \cap D_u^\epsilon = F_i^\epsilon$ and since $F_i^\epsilon \in \mathcal{F}$ and D_u^ϵ is a P-kernel of $X_\epsilon^{-1}(u)$, then by (ii) and Lem.A.5.10 we have

(iv) $\xi_\epsilon^{-1}(\gamma)^* = F_i^\epsilon \cap D_u^\epsilon = F_i^\epsilon = \xi_\epsilon^{-1}(\Lambda_i^\epsilon)$ a.s. $\forall\, i \in L_\epsilon$

where $\xi_\epsilon^{-1}(\gamma)^*$ is any upper P-envelope of $\xi_\epsilon^{-1}(\gamma)$.

Let us define $D_A^\epsilon := \cup_{u \in A} D_u^\epsilon$ for all $A \subseteq M_\epsilon$ with the usual convention $D_\emptyset^\epsilon := \emptyset$. Since D_u^ϵ is an upper P-envelope of $X_\epsilon^{-1}(u)$ and $X_\epsilon^{-1}(A) = \cup_{u \in A} X_\epsilon^{-1}(u)$, then by Lem.A.5.8 we see that D_A^ϵ is an upper P-envelope of $X_\epsilon^{-1}(A)$. Hence, if we define $\mathcal{D}_\epsilon := \{D_A^\epsilon \mid A \subseteq M_\epsilon\}$, then by (3) we have that the pavings $\mathcal{D}_1$ and $\mathcal{D}_2$ are independent. Then the complementary pavings $\mathcal{D}^1 := \{F^c \mid F \in \mathcal{D}_1\}$ and $\mathcal{D}^2 := \{F^c \mid F \in \mathcal{D}_2\}$ are independent and since the pavings $\mathcal{D}_1$ and $\mathcal{D}_2$ are stable under finite unions, then the pavings $\mathcal{D}^1$ and $\mathcal{D}^2$ are stable under finite intersections. So by [7; (2.5) p.106–107], we have that the σ-algebras $\sigma(\mathcal{D}^1)$ and $\sigma(\mathcal{D}^2)$ are independent and since $D_u^\epsilon \in \sigma(\mathcal{D}_\epsilon) = \sigma(\mathcal{D}^\epsilon)$, we see that the σ-algebras $\mathcal{F}_1$ and $\mathcal{F}_2$ are independent. So by (i) we see that ξ_1 and ξ_2 are independent measurable functions from $(\Omega, \mathcal{F})$ into $(\Gamma_1, \mathcal{G}_1)$ and $(\Gamma_2, \mathcal{G}_2)$, respectively.

Let $(\Gamma, \mathcal{G}) := (\Gamma_1 \times \Gamma_2, \mathcal{G}_1 \otimes \mathcal{G}_2)$ denote the product space and let us define $\Lambda_{ij} := \Lambda_i^1 \times \Lambda_j^2$ for all $(i,j) \in L_1 \times L_2$. Then $(\Gamma, \mathcal{G})$ is a discrete measurable space with atoms $(\Lambda_{ij} \mid (i,j) \in L_1 \times L_2)$ and $\xi(\omega) := (\xi_1(\omega), \xi_2(\omega))$ is a measurable function from $(\Omega, \mathcal{F})$ into $(\Gamma, \mathcal{G})$. Let $(i,j) \in L_1 \times L_2$ and $\gamma = (\gamma_1, \gamma_2) \in \Lambda_{ij}$ be given. Then there exist $u \in M_1$ and $v \in M_2$ such that $\gamma_1 = (i,u)$ and $\gamma_2 = (j,v)$ and so by (ii) we have

$$\xi^{-1}(\gamma) = \xi_1^{-1}(i,u) \cap \xi_2^{-1}(j,v) = F_i^1 \cap X_1^{-1}(u) \cap F_i^2 \cap X_2^{-1}(u)$$

By (3), we have that $(X_1^{-1}(u) \cap X_2^{-1}(v))^* = X_1^{-1}(u)^* \cap X_2^{-1}(v)^*$ a.s. and since $F_i^1 \cap F_j^2 \in \mathcal{F}$, then by (iv) and Lem.A.4.10 we have

$$\xi^{-1}(\gamma)^\star = F_i^1 \cap F_j^2 \cap (X_1^{-1}(u) \cap X_2^{-1}(v))^* = F_i^1 \cap F_j^2 \cap X_1^{-1}(u)^* \cap X_2^{-1}(v)^*$$
$$= \xi_1^{-1}(\gamma_1)^* \cap \xi_2^{-1}(\gamma_2)^* = \xi_1^{-1}(\Lambda_i^1) \cap \xi_2^{-1}(\Lambda_i^2) = \xi^{-1}(\Lambda_{ij}) \text{ a.s.}$$

In particular, we have $P^*(\xi = \gamma) = P_\xi(\Lambda_{ij})$ for all $\gamma \in \Lambda_{ij}$ and all $(i,j) \in L_1 \times L_2$. Since $(\Lambda_{ij} \mid (i,j) \in L_1 \times L_2\}$ are the atoms of the discrete σ-algebra $\mathcal{G}$, then by Lem.B.7.3 we conclude that ξ is strictly $(P, \mathcal{G})$-perfect and since $X_i(\omega) = \hat{X}_i(\xi_i(\omega))$ and ξ_1 and ξ_2 are independent, then by (iii) we see that X_1 and X_2 are strictly perfectly independent.

(4) $\Rightarrow$ (5): Follows from Thm.2.6.2.

(5) $\Rightarrow$ (1): Evident!

Thus, the equivalence of (1)–(5) is proved under the assumption that M_1 and M_2 are finite. So suppose that M_1 and M_2 are countable. Without loss of generality, we may assume that $M_1 = M_2 = \mathbf{N}$. By Thm.2.6.1 and Lem.2.7.2+4, we see that the implications "(1) $\Rightarrow$ (2) $\Rightarrow$ (3)" hold and the implication "(5) $\Rightarrow$ (1)" is evident. Hence, it suffices to show that (3) implies (5) when $M_1 = M_2 = \mathbf{N}$.

(3) $\Rightarrow$ (5): Suppose that (3) holds and that $M_1 = M_2 = \mathbf{N}$. Let $n \geq 1$ be a given integer and let us define $X_i^n := X_i \wedge (n+1)$ for $i = 1, 2$. Let $A, B \subseteq \mathbf{N}$ be given sets. By (3) and Lem.A.5.8, we have

$$
\begin{aligned}
&(X_1^{-1}(A) \cap X_2^{-1}(B))^* = \bigcup_{i \in A} \bigcup_{j \in B} \{X_1 = i\ ,\ X_2 = j\}^* \\
&= \bigcup_{i \in A} \bigcup_{j \in B} \{X_1 = i\}^* \cap \{X_2 = j\}^* = \bigcup_{i \in A} \{X_1 = i\}^* \cap \bigcup_{j \in B} \{X_2 = j\}^* \\
&= X_1^{-1}(A)^* \cap X_2^{-1}(B)^* \ \text{ a.s.}
\end{aligned}
$$

In particular, we see that (X_1^n, X_2^n) satisfies condition (3) and since X_1^n and X_2^n takes values in the finite set $\{1, \ldots, n+1\}$, then by the first part of the theorem proved above, we have that X_1^n, X_2^n are totally independent for all $n \geq 1$. Let $D \subseteq \mathbf{N}^2$ be a given and let us define $D_n := \{(i,j) \in D \mid i \leq n\ ,\ j \leq n\}$, $D_n^x := \{y \in \mathbf{N} \mid (x,y) \in D_n\}$ and $\psi_n(x) := P^*(X_2 \in D_n)$ for all $x \in \mathbf{N}$. Then we have $\psi(x) = P^*(X_2^n \in D_n)$ for all $x \in \mathbf{N}$ and since $\psi_n(x) = 0$ for all $x > n$, we have $\psi_n(X_1) = \psi_n(X_1^n)$. By the definition of (D_n, X_1^n, X_2^n) , we have $\{(X_1, X_2) \in D_n\} = \{(X_1^n, X_2^n) \in D_n\}$ and since X_1^n and X_2^n are totally independent, we have

$$
\begin{aligned}
P^*(\,(X_1, X_2) \in D) \geq P^*(\,(X_1, X_2) \in D_n) = P^*(\,(X_1^n, X_2^n) \in D_n) \\
\geq E^* \psi_n(X_1^n) = E^* \psi_n(X_1)
\end{aligned}
$$

for all $n \geq 1$. Since $D_n \uparrow D$, then by Lem.A.1.6 we have $\psi_n(x) \uparrow \psi(x) := P^*(X_2 \in D^x)$ for all $x \in \mathbf{N}$ where $D^x := \{y \in \mathbf{N} \mid (x,y) \in D\}$. So by Lem.A.2.7 we have $P^*((X_1, X_2) \in D) \geq E^* \psi(X_1)$. Hence, we see that X_1 and X_2 are totally independent and since X_2^n and X_1^n are totally independent, then in the same manner we see that X_2 and X_1 are totally independent; that is, X_1, X_2 are totally independent which completes the proof of the equivalence of (1), (2), (3) and (5) when M_1 and M_2 are countable.

(6): Let $A \in \mathcal{M}_P(X_1)$ and $B \in \mathcal{M}_P(X_2)$ be given. By Thm.2.6.1, we have that $X_1^{-1}(A)$ and $X_2^{-1}(B)$ are totally independent and since $X_1^{-1}(A) \in \mathcal{F}_P$ and $X_2^{-1}(B) \in \mathcal{F}_P$, then by Lem.2.7.1 we see that $X_1^{-1}(A)$ and $X_2^{-1}(B)$ are independent; that is, X_1 and X_2 are completely independent.

(7): Suppose that M_1 is not real measurable (the case where M_2 is not real measurable follows in the same manner). Since P_{X_1} is a probability measure on $(M_1, 2^{M_1})$, there exists a countable set $Q \subseteq M_1$ such that $P_{X_1}(Q) = 1$. Let $D \subseteq M_1 \times M_2$ be an arbitrary set and let us define $C := (Q \times M_2) \cap D$ and $D_x := \{y \in M_2 \mid (x,y) \in D\}$ for all $x \in M_1$. Since Q is countable and $C = \cup_{q \in Q} \{q\} \times D_q$, we see that $C \in 2^{M_1} \otimes 2^{M_2}$ and since $P_{X_1}(M_1 \setminus Q) = 0$ and $D \setminus C \subseteq (M_1 \setminus Q) \times M_2$, we see that $D \setminus C$ is a $P_{(X_1, X_2)}$-null set. Hence, we see that all subsets of $M_1 \times M_2$ are $P_{(X_1,X_2)}$-measurable. So by Lem.B.2 we have that $(X_1, X_2) : \Omega \to M_1 \times M_2$ is $(P, 2^{M_1} \otimes 2^{M_2})$-perfect and since $X_1 : \Omega \to (M_1, 2^{M_1})$ and $X_2 : \Omega \to (M_2, 2^{M_2})$ are independent random functions, we see that X_1 and X_2 are perfectly independent. $\square$

Theorem 2.9: *Let M_1 and M_2 be non-empty sets and let $X_1 : \Omega \to M_1$ and $X_2 : \Omega \to M_2$ be totally independent functions. Let $\phi_1 : M_1 \to [0, \infty]$ and $\phi_2 : M_2 :\to [0, \infty]$ be non-negative functions. Then we have*

(1) *The σ-algebras $\mathcal{L}_P(X_1)$ and $\mathcal{L}_P(X_2)$ are independent and the algebras $\mathcal{A}_P(X_1)$ and $\mathcal{A}_P(X_2)$ are totally independent*

(2) $(E^*\phi_1(X_1)) \cdot (E^*\phi_2(X_2)) = E\{\phi_1(X_1)^*\,\phi_1(X_2)^*\} = E^*\{\phi_1(X_1)\,\phi_2(X_2)\}$

Let L_1 and L_2 be finite sets and let $Y_1 : \Omega \to L_1$, $f_1 : M_1 \to L_1$, $Y_2 : \Omega \to L_2$ and $f_2 : M_2 \to L_2$ be given functions such that $Y_i^{-1}(u) \in \mathcal{A}_P(X_i)$ for $i = 1, 2$ and all $u \in L_i$. Then we have

(3) *$f_1(X_1)$ and $f_2(X_2)$ are strictly perfectly independent and Y_1 and Y_2 are totally independent*

(4) *The algebras $\mathcal{A}_P(X_1)$ and $\mathcal{A}_P(X_2)$ are totally independent*

Proof: (1): Let $\mathcal{K}_i$ denote the paving of all P-kernels of all sets $F \in \mathcal{S}_P(X_i)$ for $i = 1, 2$. Let $G_1 \in \mathcal{K}_1$ and $G_2 \in \mathcal{K}_2$ be given. Then there exist sets $A_i \subseteq M_i$ such that $G_i = X_i^{-1}(A_i)_*$ for $i = 1, 2$. Since X_1 and X_2 are totally independent, then by Thm.2.6.1 we have that $1_{X_1^{-1}(A_1)}$ and $1_{X_2^{-1}(A_2)}$ are totally independent. So by Lem.2.7.2 we see that G_1 and G_2 are independent; that is, the pavings $\mathcal{K}_1$ and $\mathcal{K}_2$ are independent and by Lem.2.1.4 we have that $\mathcal{K}_1$ and $\mathcal{K}_2$ are stable under countable intersections. Hence, by [7; (2.5) p.106–107] we conclude that the σ-algebras $\sigma(\mathcal{K}_1)$ and $\sigma(\mathcal{K}_2)$ are independent and since $\mathcal{L}_P(X_i) = \sigma(\mathcal{K}_i)$, we have proved (1).

(2): By (1) and Lem.2.1.5, we have that $\phi_1(X_1)^*$ and $\phi_2(X_2)^*$ are independent random variables. Hence, the first equality in (2) follows from Lem.A.5.1. Let us define $\phi(x, y) := \phi_1(x)\,\phi_2(y)$ and $\Phi(x) := E^*\phi(x, X_2)$ for all $x \in M_1$ and all $y \in M_2$. Since $\phi_1 \geq 0$, then by Lem.A.2.5+9 we have $\Phi(x) = \phi_1(x)\,E^*\phi_2(X_2)$ and since $\phi_1, \phi_2 \geq 0$, we have $\phi_1(X_1)^*\,\phi_1(X_2)^* \geq \phi_1(X_1)\,\phi_2(X_2)$. Since $E^*\phi_2(X_2) \geq 0$, then by Thm.2.6.1 and Lem.A.2.5+9 we have

$$E^*\{\phi_1(X_1)\,\phi_1(X_2)\} = E^*\phi(X_1, X_2) \geq E^*\Phi(X_1) = (E^*\phi_1(X_1)) \cdot (E^*\phi_2(X_2))$$
$$= E\{\phi_1(X_1)^*\,\phi_2(X_2)^*\} \geq E^*\{\phi_1(X_1)\,\phi_2(X_2)\}$$

which completes the proof of (2).

(3+4): By Thm.2.6.1, we have that $f_1(X_1)$ and $f_2(X_2)$ are totally independent and since L_1 and L_2 are finite, then by Thm.2.8 we see that $f_1(X_1)$ and $f_2(X_2)$ are strictly perfectly independent. Since$\mathcal{A}_P(X_i)$ is an algebra and $Y_i^{-1}(u) \in \mathcal{A}_P(X_i)$

for all u in the finite set L_i , we have $Y_i^{-1}(A) \in \mathcal{A}_P(X_i)$ for all $A \subseteq L_i$. Let $\mathcal{K}_i$ denote the paving of all P-kernels of all sets $F \in \mathcal{S}_P(X_i)$. Since L_i is finite and $\mathcal{A}_P(X_i)$ is the algebra generated by $\mathcal{K}_i \cup \mathcal{S}_P(X_i)$ there exist sets $F_1^i, \ldots, F_n^i \in \mathcal{K}_i$ and $F_{n+1}^i, \ldots, F_{n+m}^i \in \mathcal{S}_P(X_i)$ such that $Y_i^{-1}(A) \in \alpha(F_1^i, \ldots, F_{n+m}^i)$ for all $A \subseteq L_i$. By the definitions of $\mathcal{K}_i$ and $\mathcal{S}_P(X_i)$, there exist sets $A_1^i, \ldots, A_{n+m}^i \subseteq M_i$ such that $F_j^i = X_i^{-1}(A_j^i)_*$ for all $1 \leq j \leq n$ and $F_j^i = X_i^{-1}(A_j^i)$ a.s. for all $n < j \leq n+m$. Let us define $h_i(x) := (1_{A_1^i}(x), \ldots, 1_{A_{n+m}^i}(x))$ for all $x \in M_i$. Then h_i is a function from M_i into the finite set $\{0,1\}^{n+m}$ such that $F_1^i, \ldots, F_{m+n}^i \in \mathcal{C}_P(h_i(X_i)$ for $i = 1, 2$. Hence, by the argument above, we have that $h_1(X_1)$ and $h_2(X_2)$ are strictly perfectly independent and since $Y_i^{-1}(A) \in \alpha(F_1^i, \ldots, F_{m+n}^i)$ for all $A \subseteq L_i$, then by Thm.2.6.2 we see that Y_1 and Y_2 are totally independent. In particular, we see that the algebras $\mathcal{A}_P(X_1)$ and $\mathcal{A}_P(X_2)$ are totally independent. $\square$

Theorem 2.10: *Let M_i be a given set and let $X_i : \Omega \to M_i$ be a given function for all i in the non-empty set I . Let $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$ be a probability space and let $\eta : \tilde{\Omega} \to (\Omega, \mathcal{F})$ be a P-measurable function satisfying $\tilde{P}_\eta = P$. If we define $Y_i(\tilde{\omega}) := X_i(\eta(\tilde{\omega}))$ for all $i \in I$ and all $\tilde{\omega} \in \tilde{\Omega}$, then we have*

(1) *If η is $(\tilde{P}, \mathcal{F})$-perfect and the functions $(X_i : \Omega \to M_i \mid i \in I)$ are perfectly independent or totally independent or completely independent, then the functions $(Y_i : \tilde{\Omega} \to M_i \mid i \in I)$ are perfectly independent or totally independent or completely independent*

(2) *If η is strictly $(\tilde{P}, \mathcal{F})$-perfect and the functions $(X_i : \Omega \to M_i \mid i \in I)$ are strictly perfectly independent then the functions $(Y_i : \tilde{\Omega} \to M_i \mid i \in I)$ are strictly perfectly independent*

(3) *The notion “completely independent” is weakly consistent and the notions “independently defined”, “perfectly independent”, “strictly perfectly independent” and “totally independent” are strongly consistent*

(4) *“independently defined” $\Rightarrow$ ”strictly perfectly independent” $\Rightarrow$ ”perfectly independent” $\Rightarrow$ ”totally independent” $\Rightarrow$ ”completely independent”*

Let $\mathcal{B}_i$ be a σ-algebra on M_i for all $i \in I$, let $(M^I, \mathcal{B}^I) = (\prod_{i \in I} M_i, \otimes_{i \in I} \mathcal{B}_i)$ denote the product space and let us define $X^I(\omega) := (X_i(\omega))_{i \in I}$ for all $\omega \in \Omega$. If $(X_i : \Omega \to (M_i, \mathcal{B}_i) \mid i \in I)$ are independent random functions, then $X^I(\omega) := (Y_i(\omega))_{i \in I}$ is a P-measurable function from Ω into $(M^I, \mathcal{B}^I)$ and we have

(5) *If X^I is $(P, \mathcal{B}^I)$-perfect, then the functions $(X_i : \Omega \to M_i \mid i \in I)$ are perfectly independent*

(6) *If X^I is strictly $(P, \mathcal{B}^I)$-perfect, then the functions $(X_i : \Omega \to M_i \mid i \in I)$ are strictly perfectly independent*

Remark: Recall that the example in Def.2.2 shows that the notion "completely independent" is *not* strongly consistent.

Proof: (1+2): Suppose that the functions (X_i) are perfectly or strictly perfectly independent and let $(\Omega_i, \mathcal{F}_i, \xi_i, \hat{X}_i)$ be chosen according to Def.2.4.1+2. Then we have $Y_i(\tilde{\omega}) = \hat{X}_i(\tilde{\xi}_i())$ where $\tilde{\xi}_i(\tilde{\omega}) := \xi_i(\tilde{\eta}(\tilde{\omega}))$ for all $i \in I$ and all $\tilde{\omega} \in \tilde{\Omega}$. Since the functions $(\xi_i : \Omega \to (\Omega_i, \mathcal{F}_i) \mid i \in I)$ are independent and $\eta : \tilde{\Omega} \to (\Omega, \mathcal{F})$ is $\tilde{P}$-measurable with distribution P, we see that $(\tilde{\xi}_i : \tilde{\Omega} \to (\Omega_i, \mathcal{F}_i) \mid i \in I)$ are independent. If η and $\xi := (\xi_i)$ are perfect or strictly perfect, then by Lem.B.6.1=2 we have that $\xi \circ \eta = (\tilde{\xi}_i)$ is perfect or strictly perfect which proves (2) and the first claim in (1). Suppose that $X : \Omega \to M$ and $Y : \Omega \to L$ are totally independent. Since η is $(\tilde{P}, \mathcal{F})$-perfect with $\tilde{P}_\eta = P$, we have

$$\tilde{P}^*(Y(\eta) \in A) = \tilde{P}^*_\eta(Y \in A) = P^*(Y \in A)$$

$$\tilde{P}^*((X(\eta), Y(\eta) \in D) = \tilde{P}^*_\eta((X, Y) \in D) = P^*((X, Y) \in D)$$

for all $A \subseteq M$ and all $D \subseteq L \times M$. Hence, we see that $(X(\eta), Y(\eta))$ are totally independent but then we see that total independence of (X_i) implies total independence of (Y_i). Suppose that the functions (X_i) are completely independent. Let $\alpha \subseteq I$ be a finite non-empty set and let $A_i \in \mathcal{M}_{\tilde{P}}(Y_i)$ be given sets for all $i \in I$. By Lem.2.1.3, we have $X_i^{-1}(A_i) \in \mathcal{M}_{\tilde{P}}(\eta)$ and since η is $(\tilde{P}, \mathcal{F})$-perfect with $\tilde{P}_\eta = P$, then by Def.2.1.9 we see that $X_i^{-1}(A_i) \in \mathcal{F}_P$ or equivalently $A_i \in \mathcal{M}_P(X_i)$. In particular, we have $P(X_i^{-1}(A_i)) = P_\eta(X_i^{-1}(A_i)) = \tilde{P}(Y_i^{-1}(A_i))$ for all $i \in \alpha$ and since the functions (X_i) are completely independent, we have

$$\prod_{i \in \alpha} \tilde{P}(Y_i^{-1}(A_i)) = \prod_{i \in \alpha} P(X_i^{-1}(A_i)) = P\left(\bigcap_{i \in \alpha} X_i^{-1}(A_i)\right)$$
$$= \tilde{P}_\eta\left(\bigcap_{i \in \alpha} X_i^{-1}(A_i)\right) = \tilde{P}\left(\bigcap_{i \in \alpha} Y_i^{-1}(A_i)\right)$$

Hence, we see that the functions (Y_i) are completely independent.

(3): By Lem.2.1.3, we see that "complete independence" is weakly consistent and by the very definition of "independently defined", we see that "independently defined" is strongly consistent. By Lem.B.7.1+2 and Lem.B.9, it follows easily that "perfectly independent" and "strictly perfectly independent" are strongly consistent.

Let consider the notion "total independence". The property (2.A) is evident and the property (2.B) follows from Thm.2.6.1. So let $(X_i : \Omega \to M_i \mid i \in I)$ be totally independent functions, let $(I^\theta \mid \theta \in \Theta)$ be a family of disjoint non-empty subsets of I and let us define $M^\theta := \prod_{i \in I_\theta} M_i$ and $Y_\theta(\omega) := (X_i(\omega))_{\in I_\theta}$ for all $\theta \in \Theta$ and all $\omega \in \Omega$. Let $\alpha, \beta \subseteq \Theta$ be given disjoint sets and let us define $Y^\alpha := (Y_\theta)_{\theta \in \alpha}$ and $Y^\beta := (Y_\theta)_{\theta \in \beta}$. If we define $\mathrm{A} := \cup_{\theta \in \alpha} I_\theta$ and $\mathrm{B} := \cup_{\theta \in \beta} I_\theta$, then A and B are disjoint non-empty subsets of I such that $Y^\alpha = X^{\mathrm{A}}$ and $Y^\beta = X^{\mathrm{B}}$. Since the functions (X_i) are totally independent, we have that $Y^\alpha = X^{\mathrm{A}}$ and

$Y^\beta = X^{\mathrm{B}}$ are totally independent. Hence, we see that the functions $(Y_\theta \mid \theta \in \Theta)$ are totally independent and so we see that "total independence" satisfies (2.C); that is, "total independence" is strongly consistent.

(4): The first implication in (4) is an easy consequence of Lem.B.9 and the second implication is evident.

Let $(X_i : \Omega \to M_i \mid i \in I)$ be a given family of perfectly independent functions and let $(\Omega_i, \mathcal{F}_i, \xi_i, \hat{X}_i)$ be chosen according to Def.2.4.1+2. If $\alpha \subseteq I$ is a non-empty set, we let $M^\alpha := \prod_{i \in \alpha} M_i$ and $(\Omega^\alpha, \mathcal{F}^\alpha) := (\prod_{i \in \alpha} \Omega_i, \times_{i \in \alpha} \mathcal{F}_i)$ the product spaces and we define $\xi_\alpha := (\xi_i)_{i \in \alpha}$, $\hat{X}^\alpha := (\hat{X}_i)_{i \in \alpha}$ and $X^\alpha := (X_i)_{i \in \alpha}$. By Def.2.4.1+2, we have that ξ_I is $(P, \mathcal{F}^I)$-perfect with distribution $\times_{i \in \alpha} P_{\xi_i}$. So by Lem.B.6.1 and Lem.B.7.1 we see that ξ_α is $(P, \mathcal{F}^\alpha)$ perfect. Let $\alpha, \beta \subseteq I$ be disjoint non-empty sets and let $D \subseteq M^\alpha \times M^\beta$ be given. Let us define $\gamma := \alpha \cup \beta$ and $f(x) := P^*(X^\beta \in D_x)$ for all $x \in M^\alpha$ where $D_x := \{y \in M^\beta \mid (x, y) \in D\}$. Then we have $D \subseteq M^\alpha \times M^\beta = M^\gamma$ and $X^\gamma = (X^\alpha, X^\beta) = \hat{X}^\gamma \circ \xi_\gamma$. Hence, if we define $C := (\hat{X}^\gamma)^{-1}(D)$ and $C_\upsilon := \{\theta \in \Omega^\beta \mid (\upsilon, \theta) \in C\}$ for all $\upsilon \in \Omega^\alpha$, then we have $\{(X^\alpha, X^\beta) \in D\} = \xi_\gamma^{-1}(C)$ and $\xi_\beta^{-1}(C_\upsilon) = \{X^\beta \in D_{\xi_\alpha(\upsilon)}\}$ for all $\upsilon \in \Omega^\alpha$. Since ξ_β is $(P, \mathcal{F}^\beta)$-perfect, we have $f(\xi_\alpha(\upsilon)) = P^*(X^\beta \in D_{\xi_\alpha(\upsilon)}) = P^*_{\xi_\beta}(C_\upsilon)$ and since ξ_α and ξ_γ are perfect and $P_{\xi_\gamma} = P_{\xi_\alpha} \times P_{\xi_\beta}$, then by Thm.A.6.1 we have

$$P^*(\,(X^\alpha, X^\beta) \in D) = P^*(\,X^\gamma \in D) = P^*_{\xi_\gamma}(C) \geq \int^* P^*_{\xi_\beta}(C_\upsilon)\, P_{\xi_\alpha}(d\upsilon)$$

$$= \int^* f(\xi_\alpha(\upsilon))\, P_{\xi_\alpha}(d\upsilon) = E^* f(X^\alpha)$$

Hence, we see that X^α and X^β are totally independent; that is the functions (X_i) are totally independent which proves the third implication in (4).

Suppose that $(X_i : \Omega \to M_i \mid i \in I)$ are totally independent functions and let us define $M^\alpha := \prod_{i \in \alpha} M_i$ and $\mathcal{B}^\alpha := \otimes_{i \in \gamma} \mathcal{M}_P(X_i)$ for all non-empty sets $\alpha \subseteq I$. Since $X_i : \Omega \to (M_i, \mathcal{M}_P(X_i)$ is P-measurable, we see that $X^\alpha : \Omega \to (M^\alpha, \mathcal{B}^\alpha)$ is P-measurable and so we have $\mathcal{B}^\alpha \subseteq \mathcal{M}_P(X^\alpha)$ for all non-empty sets $\alpha \subseteq I$. Let $\alpha, \beta \subseteq I$ be disjoint non-empty sets. By (3), we have that $X^\alpha : \Omega \to M^\alpha$ and $X^\beta : \Omega \to M^\beta$ are totally independent. So by Lem.2.1.1 and Thm.2.9.1 we see that $X^\alpha : \Omega \to (M^\alpha, \mathcal{M}_P(X^\alpha))$ and $X^\beta :\to (M^\beta, \mathcal{M}_P(X^\beta))$ are independent random function and since $\mathcal{B}^\gamma \subseteq \mathcal{M}_P(X^\gamma)$, we conclude that $X^\alpha : \Omega \to (M^\alpha, \mathcal{B}^\alpha)$ and $X^\beta : \Omega \to (M^\beta, \mathcal{B}^\beta)$ are independent random functions whenever α and β are non-empty disjoint subsets of I . But then a standard argument shows that the functions $(X_i : \Omega \to (M_i, \mathcal{M}_P(X_i)) \mid i \in I)$ are independent; that is, the functions (X_i) are completely independent which proves the last implication in (4).

(5) and (6) are immediate consequences of the definitions. □

Theorem 2.11: *Let $(L, \mathcal{A})$ be a measurable space and let $\xi : \Omega \to (L, \mathcal{A})$ be discrete random function with discrete range Q . Let M be a given set and let $X : \Omega \to M$ and $\psi : L \times M \to \bar{\mathbf{R}}$ be given functions such that $\psi(q, x) \geq 0$ for all $(q, x) \in Q \times M$.*

Let C^* denote a fixed P-hull of C whenever $C \subseteq \Omega$ and let $\psi(q,X)^*$ denote an upper P-envelope of $\psi(q,X)$ for all $q \in Q$ and let us define $\psi^\circ(u) := E^*\psi(u,X)$ for all $u \in L$. Then $\psi^\circ \in L(P_\xi)$ and we have

(1) *If $\{\xi = q\}$ and $\psi(q,X)^*$ are independent for all $q \in Q$, then ξ and X are upper ψ-independent and we have*

(a) $E^*\psi(\xi,X) = E\psi^\circ(\xi) = \sum_{q\in Q} P(\xi = q)\, E^*\psi(q,X)$

(2) *ξ and X are totally independent if and only if $\{\xi = q\}$ and $\{X \in A\}^*$ are independent for all $q \in Q$ and all $A \subseteq M$ and if so, then (1.a) holds for every function $\psi : L \times M \to \bar{\mathbf{R}}$ satisfying $\psi(q,x) \geq 0$ for all $(q,x) \in Q \times M$*

Proof: (1): Since $P(\xi \in Q) = 1$ and $\psi(q,x) \geq 0$ for all $(q,x) \in Q \times M$, we have $\psi^\circ \geq 0$ P_ξ-a.s. and

$$\psi(\xi,X) = \sum_{q\in Q} 1_{\{\xi=q\}}\, \psi(q,X) = \sup_{q\in Q} \left\{ 1_{\{\xi=q\}}\, \psi(q,X) \right\} \text{ a.s.}$$

So by Def.A.9.1, we see that $\psi^\circ \in L(P_{X_1})$ and since Q is countable and $\{\xi = q\} \in \mathcal{F}_P$ for all $q \in Q$, then by Lem.A.4.3+10 we have

$$\psi(\xi,X)^* = \sup_{q\in Q} \left\{ 1_{\{\xi=q\}}\, \psi(q,X)^* \right\} = \sum_{q\in Q} 1_{\{\xi=q\}}\, \psi(q,X)^* \text{ a.s.}$$

Since $\{\xi = q\}$ and $\psi(q,X)^*$ are ordinary independent and $\psi(q,X)^* \geq 0$ a.s. then by Lem.A.5.1 and Beppo-Levi's theorem we have

$$E^*\psi(\xi,X) = E\psi(\xi,X)^* = \sum_{q\in Q} E\Big\{ 1_{\{\xi=q\}}\, \psi(q,X)^* \Big\} = \sum_{q\in Q} P(\xi = q)\, E\psi(q,X)^*$$

$$= \sum_{q\in Q} P(\xi = q)\, E^*\psi(q,X) = \sum_{q\in Q} P(\xi = q)\, \psi^\circ(q) = E\psi^\circ(\xi)$$

which proves (1.a).

(2): Suppose that ξ and X are totally independent and let $q \in Q$ and $A \subseteq M$ be given. Since $\{\xi = q\} \in \mathcal{F}_P$, then by Lem.2.1.4 we have $\{\xi = q\} \in \mathcal{L}_P(\xi)$ and by Lem.2.1.5, we have $\{X \in A\}^* \in \mathcal{L}_P(X)$. So by Thm.2.9.1 we conclude that $\{\xi = q\}$ and $\{X \in A\}^*$ are independent. Suppose that $\{\xi = q\}$ and $\{X \in A\}^*$ are independent for all $q \in Q$ and all $A \subseteq M$ and let $\mathcal{K}$ denote the paving of all P-kernels of all sets in $\mathcal{S}_P(X)$. By Lem.A.5.6, we see that every set in $K \in \mathcal{K}$ is of the form $K = \Omega \setminus \{X \in A\}^*$ a.s. Hence, we see that $\{\xi = q\}$ and K are independent for all $K \in \mathcal{K}$ and by Lem.2.1.4 we have that $\mathcal{K}$ is stable under countable intersections. Since $\mathcal{L}_P(X) = \sigma(\mathcal{K})$, then by [7; (2.5) p.106–107] we conclude that $\{\xi = q\}$ and $\mathcal{L}_P(X)$ are independent for all $q \in Q$. So by Lem.2.1.5 we see that $\{\xi = q\}$ and $f(X)^*$ are independent for every function $f : M \to \bar{\mathbf{R}}$. Thus, by (1) we see that (1.a) holds for every function $\psi : L \times M \to \bar{\mathbf{R}}$ satisfying $\psi(q,x) \geq 0$ for all $(q,x) \in Q \times M$. In particular, we see that ξ and X are totally independent. □

3. Symmetrization It is well-known that symmetrization is a useful tool for establishing stochastic inequalities and in this section we shall study various ways of symmetrizing functions taking values in a vector space.

Let Λ is a vector space over the reals. Then we let Λ^* denote the (algebraic) *dual space*; that is, the set of all linear functionals from Λ into $\mathbf{R}$. If $\mathcal{L}$ is a σ-algebra on Λ, we say that $(\Lambda,\mathcal{L})$ is a *measurable linear space* if the addition $(u,v) \curvearrowright u+v$ is measurable from $(\Lambda\times\Lambda,\mathcal{L}\otimes\mathcal{L})$ into $(\Lambda,\mathcal{L})$ and the scalar multiplication $(\alpha,u) \curvearrowright \alpha u$ is measurable from $(\mathbf{R}\times\Lambda,\mathcal{B}(\mathbf{R})\otimes\mathcal{L})$ into $(\Lambda,\mathcal{L})$. Recall that we have

(3.1) If $\Xi \subseteq \Lambda^*$ is a set of linear functionals, then $(\Lambda,\sigma(\Xi))$ is a measurable linear space where $\sigma(\Xi)$ denotes the smallest σ-algebra on Λ making all functions in Ξ measurable

(3.2) If $(\Lambda_i,\mathcal{L}_i)$ is a measurable linear space for all i in the non-empty set I, then the product space $(\prod_{i\in I}\Lambda_i, \otimes_{i\in I}\mathcal{L}_i)$ is a measurable linear space with respect to the usual (coordinatewise) addition and scalar multiplication

(3.3) If Λ is a topological linear space such that Λ is hereditarily Lindelöf, then $(\Lambda,\mathcal{B}(\Lambda))$ is a measurable linear space where $\mathcal{B}(\Lambda)$ denotes the Borel σ-algebra on Λ

In particular, we see that $(\mathbf{R}^T,\mathcal{B}^T)$ is a measurable linear space and recall that a *stochastic process with parameter set* T is a random function from Ω into $(\mathbf{R}^T,\mathcal{B}^T)$. Recall that a function $\phi:\Lambda\to\bar{\mathbf{R}}$ is *subadditive* if $\phi(x+y)\le\phi(x)\dot{+}\phi(y)$ for all $x,y\in\Lambda$ where we use the upper convention for addition on the extended real line (see Appendix A) and recall that a *seminorm* is a subadditive function $\phi:\Lambda\to[0,\infty]$ satisfying $\phi(\alpha\, x)=|\alpha|\,\phi(x)$ for all $x\in\Lambda$ and all $\alpha\in\mathbf{R}$. Observe that an arbitrary supremum of subadditive functions is subadditive and that an arbitrary supremum of seminorms is a seminorm.

Defnition 3.1: (Symmetric functions) Let $n\ge 1$ be a given integer. Then we let $\mathrm{E}^n := \{-1,1\}^n$ denote the set of n-vectors of ± 1's. Note that E^n is a group under the multiplication $\eta\delta := (\eta_1\cdot\delta_1,\dots,\eta_n\cdot\delta_n)$ for $\eta=(\eta_1,\dots,\eta_n)\in\mathrm{E}^n$ and $\delta=(\delta_1,\dots,\delta_n)\in\mathrm{E}^n$. We let $\mathcal{E}^n$ denote the set of all subsets of E^n and we let B^n denote the uniform probability measure on $(\mathrm{E}^n,\mathcal{E}^n)$; that is, the probability measure given by $\mathrm{B}^n(\{\delta\}) := 2^{-n}$ for all $\delta\in\mathrm{E}^n$. Recall that a *Bernoulli vector* is a random function $\varepsilon:\Omega\to(\mathrm{E}^n,\mathcal{E}^n)$ such that $P_\varepsilon=\mathrm{B}^n$ or equivalently a random function $\varepsilon:\Omega\to(\mathrm{E}^n,\mathcal{E}^n)$ such that the coordinate functions $\varepsilon_1,\dots,\varepsilon_n$ are independent and takes the values $+1$ and -1 with probability $\frac{1}{2}$.

Let Λ be a vector space and let us define $T_\delta(x) := (\delta_1\,x_1,\dots,\delta_n\,x_n)$ for all $x=(x_1,\dots,x_n)\in\Lambda^n$ and all $\delta=(\delta_1,\dots,\delta_n)\in\mathrm{E}^n$. Then $T_\delta:\Lambda^n\to\Lambda^n$ is a bijection satisfying (see Lem.B.6.8)

(1) $T_\delta^{-1} = T_\delta$ and $T_\delta\circ T_\eta = T_\eta\circ T_\delta = T_{\eta\delta} \quad \forall\,\delta,\eta\in\mathrm{E}^n$

(2) If $(\Lambda, \mathcal{L})$ is a measurable linear space and $(\Lambda^n, \mathcal{L}^n)$ denotes the n-fold product space, then $T_\delta : (\Lambda^n, \mathcal{L}^n) \to (\Lambda^n, \mathcal{L}^n)$ is a measurable bijection and T_δ is strictly $(\mu, \mathcal{L}^n)$-perfect for every probability measure μ on $(\Lambda^n, \mathcal{L}^n)$

Let $(\Lambda, \mathcal{L})$ be a measurable linear space and let $X_1, \ldots, X_n : \Omega \to (\Lambda, \mathcal{L})$ be random functions. Recall, that the random vector $X = (X_1, \ldots, X_n)$ is *$\mathcal{L}$-symmetric* if and only if X and $T_\delta(X)$ have the same distribution on $(\Lambda^n, \mathcal{L}^n)$ for all $\delta_1, \ldots, \delta_n \in \mathrm{E}^n$. If $X_1, \ldots, X_n : \Omega \to \Lambda$ are arbitrary functions, we say that the random vector $X = (X_1, \ldots, X_n)$ is *totally symmetric* if X and $T_\delta(X)$ have the same outer distribution for all $\delta \in \mathrm{E}^n$.

In [4; p.289], R.M. Dudley defines a “symmetric vector function” to be a function $X = (X_1, \ldots, X_n) : \Omega \to \Lambda^n$ such that $X_1, \ldots, X_n$ are independently defined functions and the probability spaces $(\Omega_i, \mathcal{F}_i, P_i)$ and the functions $\hat{X}_i$ from Def.2.3.1+2 are of the form

$$(\Omega_i, \mathcal{F}_i, P_i) = (\Omega_i' \times \Omega_i', \mathcal{F}_i' \otimes \mathcal{F}_i', P_i' \otimes P_i') \quad \forall\, i = 1, \ldots, n$$

$$\hat{X}_i(\omega_1, \omega_2) = Y_i(\omega_1) - Y_i(\omega_2) \quad \forall\, (\omega_1, \omega_2) \in \Omega_i' \times \Omega_i' \;\forall\, i = 1, \ldots, n$$

that is, if X is the standard symmetrization of some independently defined functions $Y_1, \ldots, Y_n$ (see Def.3.2.2). In Thm.3.4.5 below, we shall that the standard symmetrization of strictly perfectly independent functions $X_1, \ldots, X_n : \Omega \to \Lambda$ is totally symmetric. So by Thm.2.10.4, we see that Dudley’s concept of “symmetry” implies “total symmetry”.

It is tempting to believe that if $X_1, X_2 : \Omega \to (\Lambda, \mathcal{L})$ independently defined $\mathcal{L}$-symmetric random functions with the same outer distribution, then (X_1, X_2) is totally symmetric. The following example shows that this is *not* the case and the example shows that “total symmetry” is a rather restrictive assumption.

Example: Let $(\Upsilon, \mathcal{G}, Q)$ be a non-atomic probability space such that $\{\upsilon\} \in \mathcal{G}$ for all $\upsilon \in \Upsilon$ and let $(\Omega, \mathcal{F}.P) := (\Upsilon \times \Upsilon, \mathcal{G} \otimes \mathcal{G}, Q \otimes Q)$ denote the product space. Let $\xi : \Upsilon \to \mathbf{R}$ be a symmetric $\mathcal{G}$-measurable random variable on $(\Upsilon, \mathcal{G}, Q)$ and let $f : \Upsilon \to \mathbf{R}$ be an arbitrary function, Let us define the stochastic processes $X_1, X_2 : \Omega \to \mathbf{R}^\Upsilon$ as follows

$$X_i(t, \omega_1, \omega_2) := \begin{cases} \xi(\omega_i) & \text{if } t \in \Upsilon \text{ , } (\omega_1, \omega_2) \in \Omega \text{ and } t \neq \omega_i \\ f(\omega_i) & \text{if } t \in \Upsilon \text{ , } (\omega_1, \omega_2) \in \Omega \text{ and } t = \omega_i \end{cases}$$

for $i = 1, 2$. Let $|| \cdot ||_\Upsilon$ denote the sup-norm on $\mathbf{R}^\Upsilon$. Then I claim that we have

(3) $X_1, X_2 : \Omega \to \mathbf{R}^\Upsilon$ are independently defined $\mathcal{B}^\Upsilon$-measurable functions such that (X_1, X_2) is $\mathcal{B}^\Upsilon$-symmetric and X_1 and X_2 have the same outer distribution

(4) If f is Q-measurable, then the sup-norms $||X_1||_\Upsilon$, $||X_2||_\Upsilon$, $||X_1 + X_2||_\Psi$ and $||X_1 - X_2||_\Upsilon$ are P-measurable.

(5) If $f \equiv 1$ and ξ only takes the values ± 1 with $P(\xi = 1) = P(\xi = -1) = \frac{1}{2}$, then (X_1, X_2) is *not* totally symmetric and we have

  (a) $\|X_1(\omega) + X_2(\omega)\|_{||\Upsilon} = 2 \quad \forall\, \omega \in \Omega$

(b) $\quad P(\|X_1 - X_2\|_\Upsilon = 0) = \frac{1}{4} \quad , \quad P(\|X_1 - X_2\|_\Upsilon = 2) = \frac{3}{4}$

Proof of (3): Since $\mathcal{G}$ contains all singletons, we see that $X_i(t,\cdot)$ is $\mathcal{G}$-measurable for all $t \in \Upsilon$; that is, X_1 and X_2 are $\mathcal{B}^\Upsilon$-measurable functions from Ω into $\mathbf{R}^\Upsilon$. Since Q is non-atomic and ξ is symmetric, we see that (X_1, X_2) is $\mathcal{B}^\Upsilon$-symmetric and by Thm.A.8 and the definition of X_1 and X_2 , we see that X_1 and X_2 are independently defined functions with the same outer distribution.

Proof of (4+5): Let $\omega = (\omega_1, \omega_2) \in \Omega$ be given. By the definition of X_1 and X_2 , we have

$$\|X_1(\omega)\|_\Upsilon = |\xi(\omega_1)| \vee |f(\omega_1)| \quad , \quad \|X_2(\omega)\|_\Upsilon = |\xi(\omega_2)| \vee |f(\omega_2)|$$

$$\|X_1(\omega) + X_2(\omega)\|_T = |\xi(\omega_1) + \xi(\omega_2)| \vee |\xi(\omega_1) + f(\omega_2)| \vee |f(\omega_1) + \xi(\omega_2)|$$

$$\|X_1(\omega) - X_2(\omega)\|_T = |\xi(\omega_1) - \xi(\omega_2)| \vee |\xi(\omega_1) - f(\omega_2)| \vee |f(\omega_1) - \xi(\omega_2)|$$

In particular, we see that (4) holds. Suppose that $f \equiv 1$ and that ξ only takes the values ± 1 with $P(\xi = 1) = P(\xi = -1) = \frac{1}{2}$. By the second equality derived above, we see that $||X_1(\omega) + X_2(\omega)||_\Upsilon = 2$ for all $\omega \in \Omega$. By the last equality derived above, we see that $||X_1(\omega) - X_2(\omega)||_\Upsilon = 0$ if $(\xi(\omega_1), \xi(\omega_2) = (1,1)$ and that $||X_1(\omega) - X_2(\omega)||_\Upsilon = 2$ if $(\xi(\omega_1), \xi(\omega_2) \neq (1,1)$ which proves (5.b). Since $||X_1 + X_2||_\Upsilon$ and $||X_1 - X_2||_\Upsilon$ do not have the same distribution, we see that (X_1, X_2) is not totally symmetric.

Definition 3.2: (Symmetrizations) Let Λ be a vector space and let $X_1, \ldots, X_n : \Omega \to \Lambda$ be given functions. Then there are two methods of symmetrizing the vector function $X = (X_1, \ldots, X_n)$. Let $\epsilon_i(\delta) := \delta_i$ for $\delta = (\delta_1, \ldots, \delta_n) \in \mathrm{E}^n$ denote the projection maps for $i = 1, \ldots, n$. Then $(\epsilon_1, \ldots, \epsilon_n)$ is a Bernoulli vector on the probability space $(\mathrm{E}^n, \mathcal{E}^n, \mathrm{B}^n)$ and we define *the Bernoulli symmetrization* $X^\epsilon = (X_1^\epsilon, \ldots, X_n^\epsilon)$ to be the function $X^\epsilon : \mathrm{E}^n \times \Omega \to \Lambda^n$ given by

(1) $\quad X_i^\epsilon(\delta, \omega) := \epsilon_i(\delta)\, X_i(\omega) \quad \forall\, (\delta, \omega) \in \mathrm{E}^n \times \Omega \;\forall\, i = 1, \ldots, n$

Similarly, we define *the standard symmetrization* $X^s := (X_1^s, \ldots, X_n^s)$ to be the function $X^s : \Omega \times \Omega \to \Lambda^n$ given by

(2) $\quad X_i^s(\omega_1, \omega_2) = X_i(\omega_1) - X_i(\omega_2) \quad \forall\, (\omega_1, \omega_2) \in \Omega \times \Omega \;\forall\, i = 1, \ldots, n$

Below we shall see that X^ϵ is totally symmetric on the product probability space $(\mathrm{E}^n \times \Omega, \mathcal{E}^n \otimes \mathcal{F}, \mathrm{B}^n \otimes P)$. Since $I(\omega_1, \omega_2) := (\omega_2, \omega_1)$ is a measurable bijection of $(\Omega \times \Omega, \mathcal{F} \otimes \mathcal{F})$ onto itself with $I^{-1} = I$, then by Lem.B.6.8 we have that X^s and $-X^s$ have the same outer distribution on the probability space $(\Omega \times \Omega, \mathcal{F} \otimes \mathcal{F}, P \otimes P)$. Below we shall see that X^s is totally symmetric on the probability space $(\Omega \times \Omega, \mathcal{F} \otimes \mathcal{F}, P \otimes P)$ provided that $X_1, \ldots, X_n$ are strictly perfectly independent but I don't know if X^s is totally symmetric if $X_1, \ldots, X_n$ are perfectly independent.

Let $X_1,\dots,X_n:\Omega\to\Lambda$ be given function and let $X^s=(X_1^s,\dots,X_n^s)$ denote the standard symmetrization of $X=(X_1,\dots,X_n)$. Let $A,B\subseteq\Lambda^n$ be given sets and let us define $A+B:=\{x+y\mid x\in A\ ,\ y\in B\}$. Since

$$\{\,T_\delta(X^s)\in A\}\subseteq(\Omega\times\{T_\delta(X)\notin B\})\cup(\{T_\delta(X)\in A+B\}\times\Omega)$$

for all $\delta\in\mathrm{E}^n$, then by Thm.A.8 we have

(3) $\quad (P\otimes P)^*(T_\delta(X^s)\in A)\le P^*(T_\delta(X)\notin B)+P^*(T_\delta(X)\in A+B)$

for all $\delta\in\mathrm{E}^n$ and all $A,B\subseteq\Lambda^n$. Note that the example given in Def.3.1 shows that total symmetry is an unpleasant assumption. Below we shall that we may bypass total symmetry by introducing the following paving of X-*symmetric* sets. Let $X_1,\dots,X_n:\Omega\to\Lambda$ be given functions and let $X^s=(X_1^s,\dots,X_n^s)$ denote the standard symmetrization of $X:=(X_1,\dots,X_n)$. Then we let $\Sigma_P(X)$ denote the paving of all sets $A\subseteq\Lambda^n$ satisfying

(4) $\quad (P\otimes P)^*(X^s\in A)\le P^*(T_\delta(X)\notin B)+P^*(T_\delta(X)\in A+B)\ \ \forall\,B\subseteq\Lambda^n\ \forall\,\delta\in\mathrm{E}^n$

Lemma 3.3: *Let $(\Lambda,\mathcal{L})$ be a measurable linear space and let $(\Lambda^n,\mathcal{L}^n)$ denote the n-fold product space. Let $X_1,\dots,X_n:\Omega\to\Lambda$ be given functions and let $X^\epsilon=(X_1^\epsilon,\dots,X_n^\epsilon)$ and $X^s=(X_1^s,\dots,X_n^s)$ denote the Bernoulli and standard symmetrizations of $X:=(X_1,\dots,X_n)$. If we define $(\Omega_2,\mathcal{F}_2,P_2):=(\Omega\times\Omega,\mathcal{F}\otimes\mathcal{F},P\otimes P)$ and $(\tilde\Omega,\tilde{\mathcal{F}},\tilde P):=(\mathrm{E}^n\times\Omega,\mathcal{E}^n\otimes\mathcal{F},\mathrm{B}^n\otimes P)$, then we have*

(1) *If $X_1,\dots,X_n$ are independently defined or perfectly independent or strictly independent or totally independent, then so are $X_1^\epsilon,\dots,X_n^\epsilon$ on the probability space $(\tilde\Omega,\tilde{\mathcal{F}},\tilde P)$*

(2) *If $X_1,\dots,X_n$ are independently defined or perfectly independent or strictly perfectly independent, then so are $X_1^s,\dots,X_n^s$ on the probability space $(\Omega_2,\mathcal{F}_2,P_2)$*

(3) *If $X_1,\dots,X_n:\Omega\to(\Lambda,\mathcal{L})$ are random functions such that X is $\mathcal{L}$-symmetric and $(P,\mathcal{L}^n)$-perfect, then X is totally symmetric*

(4) *If $X_1,\dots,X_n:\Omega\to(\Lambda,\mathcal{L})$ are independent random functions such that X_i and $-X_i$ have the same distribution on $(\Lambda,\mathcal{L})$, then X is $\mathcal{L}$-symmetric*

(5) *If $X_1,\dots,X_n:\Omega\to(\Lambda,\mathcal{L})$ are independent random functions, then $X_1^s,\dots,X_n^s:\Omega\times\Omega\to(\Lambda,\mathcal{L})$ are independent random functions and X^s is $\mathcal{L}$-symmetric*

Proof: (1+2): The “independently defined” case is evident. The “perfectly independent” and “strictly independent” cases are easy consequences of Lem.B.6.1+2 and Lem.B.7.1. The “totally independent” case for the Bernoulli symmetrization follows easily from Thm.2.11.

(3): Let $\delta \in \mathrm{E}^n$ be given. Since X is $(P, \mathcal{L}^n)$-perfect, then by Def.3.3.2 and Lem.B.6.1 we have that $T_\delta(X)$ is $(P, \mathcal{L}^n)$-perfect and since X is $\mathcal{L}$-symmetric, we have that $P_X = P_{T_\delta(X)}$. Hence, by Lem.B.1.2 we see that X and $T_\delta(X)$ have the same outer distribution for all $\delta \in \mathrm{E}^n$; that is, X is totally symmetric.

(4): Evident!

(5): Let us define $Y_i(\omega_1, \omega_2) := X_i(\omega_1)$ and $Z_i(\omega_1, \omega_2) := X_i(\omega_2)$ for all $(\omega_1, \omega_2) \in \Omega \times \Omega$ and all $i = 1, \dots, n$. Then $Y_1, \dots, Y_n, Z_1, \dots, Z_n : \Omega \times \Omega \to (\Lambda, \mathcal{L})$ are independent random function on the probability space $(\Omega_2, \mathcal{F}_2, P_2)$ such that $Y_i - Z_i$ and $Z_i - Y_i$ have the same distribution on $(\Lambda, \mathcal{L})$ for all $i = 1, \dots, n$ and since $X_i^s = Y_i - Z_i$, we see that (5) follows from (4). □

Theorem 3.4: *Let Λ be a vector space, let $X_1, \dots, X_n : \Omega \to \Lambda$ be given functions and let $\varepsilon = (\varepsilon_1, \dots, \varepsilon_n)$ be a random function from Ω into $(\mathrm{E}^n, \mathcal{E}^n)$ such that ε and $X = (X_1, \dots, X_n)$ are totally independent. Let $X^\epsilon = (X_1^\epsilon, \dots, X_n^\epsilon)$ and $X^s = (X_1^s, \dots, X_n^s)$ denote the Bernoulli and standard symmetrizations of X and let us define $X_\varepsilon := (\varepsilon_1 X_1, \dots, \varepsilon_n X_n)$. If we define $(\Omega_2, \mathcal{F}_2, P_2) := (\Omega \times \Omega, \mathcal{F} \otimes \mathcal{F}, P \otimes P)$ and $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P}) := (\mathrm{E}^n, \mathcal{E}^n \otimes \mathcal{F}, \mathrm{B}^n \otimes P)$, then we have*

(1) $P_2^*(X^s \in A) \le \tilde{P}^*(X^\epsilon \notin B) + \tilde{P}^*(X^\epsilon \in A + B) \;\; \forall A \in \Sigma_P(X) \; \forall B \subseteq \Lambda^n$

(2) $P_2^*(X^s \in A) \le P^*(X_\varepsilon \notin B) + P^*(X_\varepsilon \in A + B) \;\; \forall A \in \Sigma_P(X) \; \forall B \subseteq \Lambda^n$

(3) *X^ϵ is totally symmetric on the probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$ and if ε is $\mathcal{E}^n$-symmetric, then X_ε is totally symmetric*

(4) *If X^s is totally symmetric, then $A \in \Sigma_P(X)$ for all $A \subseteq \Lambda^n$*

(5) *If $X_1, \dots, X_n$ are strictly perfectly independent, then X^s is totally symmetric on the probability space $(\Omega_2, \mathcal{F}_2, P_2)$*

(6) *If $X_1, \dots, X_n$ are perfectly independent, then $A \in \Sigma_P(X)$ for all $A \subseteq \Lambda^n$*

Let $\mathcal{L}$ be a given σ-algebra on Λ such that $(\Lambda, \mathcal{L})$ is a measurable linear space and $X_1, \dots, X_n : \Omega \to (\Lambda, \mathcal{L})$ are random functions. If $(\Lambda^n, \mathcal{L}^n)$ denote the n-fold product space, then $X : \Omega \to (\Lambda^n, \mathcal{L}^n)$ and $X^s : \Omega_2 \to (\Lambda^n, \mathcal{L}^n)$ are random functions and we have

(7) *If X^s is $\mathcal{L}$-symmetric, then $\mathcal{L}^n \subseteq \Sigma_P(X)$ and we have*

(a) $P_2(X^s \in A) \le P_*(X_\varepsilon \notin B) + P_*(X_\varepsilon \in A + B) \;\; \forall A \in \mathcal{L}^n \; B \subseteq \Lambda^n$

(b) $P_2(X^s \in A) \le \tilde{P}_*(X^\epsilon \notin B) + \tilde{P}_*(X^\epsilon \in A + B) \;\; \forall A \in \mathcal{L}^n \; \forall B \subseteq \Lambda^n$

Proof: (1): Let us define $\epsilon_i(\delta, \omega) := \delta_i$ and $Y_i(\delta, \omega) := X_i(\omega)$ for all $\delta = (\delta_1, \dots, \delta_n) \in \mathrm{E}^n$ and all $\omega \in \Omega$. Then $\epsilon = (\epsilon_1, \dots, \epsilon_n)$ and $Y = (Y_1, \dots, Y_n)$ are independently defined function. So by Thm.2.10.4, Thm.2.11 and Thm.A.8, we have

$$\tilde{P}^*(X^\epsilon \in D) = 2^{-n} \sum_{\delta \in \mathrm{E}^n} \tilde{P}^*(T_\delta(Y) \in D) = 2^{-n} \sum_{\delta \in \mathrm{E}^n} P^*(T_\delta(X) \in D)$$

for all $D \subseteq \Lambda^n$. Hence, we see that (1) follows from the definition of $\Sigma_P(X)$.

(2): Let $A \in \Sigma_P(X)$ and $B \subseteq \Lambda^n$ be given. Since ε and X are totally independent and $X_\varepsilon = T_\varepsilon(X)$, then by Def.3.3.4 and Thm.2.11 we have

$$\begin{aligned} P_2^*(X^s \in A) &\leq \sum_{\delta \in \mathrm{E}^n} P(\varepsilon = \delta) \{P^*(T_\delta(X) \notin B) + P^*(T_\delta(X) \in A + B)\} \\ &= P^*(X_\varepsilon \notin B) + P^*(X_\varepsilon \in A + B) \end{aligned}$$

which proves (2).

(3): Suppose that ε is $\mathcal{E}^n$-symmetric and let $B \subseteq \Lambda^n$ and $\delta \in \mathrm{E}^n$ be given. Since ε and X are totally independent, then by Thm.2.6.1 we have that ε and $T_\delta(X)$ are totally independent and since ε is symmetric, we have that $P(\varepsilon = \delta) = P(\varepsilon = \eta\delta)$ for all $\eta \in \mathrm{E}^n$. Since $\eta \curvearrowright \eta\delta$ is a bijection of E^n onto E^n , then by Def.3.3.1 and Thm.2.11 we have

$$\begin{aligned} P^*(X_\varepsilon \in B) &= \sum_{\eta \in \mathrm{E}^n} P(\varepsilon = \eta)\, P^*(T_\eta(X) \in B) = \sum_{\eta \in \mathrm{E}^n} P(\varepsilon = \eta\delta)\, P^*(T_{\eta\delta}(X) \in B) \\ &= \sum_{\eta \in \mathrm{E}^n} P(\varepsilon = \eta)\, P^*(T_\delta(T_\eta(X)) \in B) = P^*(T_\delta(X_\varepsilon) \in B) \end{aligned}$$

Hence, we see that X_ε and $T_\delta(X_\varepsilon)$ have the same outer distribution for all $\delta \in \mathrm{E}^n$; that is, X_ε is totally symmetric. If we define $\epsilon = (\epsilon_1, \ldots, \epsilon_n)$ and $Y = (Y_1, \ldots, Y_n)$ as in the proof of (1), we have $X^\epsilon = (\epsilon_1 Y_1, \ldots, \epsilon_n Y_n)$ and since ϵ and Y are independently defined, then by Thm.2.10.4 we have that ϵ and Y are totally independent. Since ϵ is totally symmetric, then by the argument above we see that X^ϵ is totally symmetric on the probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$.

(4): Immediate consequence of Def.3.3.3.

(5+6): Suppose that $X_1, \ldots, X_n$ are perfectly independent and let $(\Omega_i, \mathcal{F}_i, \xi_i, \hat{X}_i)$ be chosen according Def.2.4.1+2. Let $(\Omega^n, \mathcal{F}^n) := (\prod_{i=1}^n \Omega_i, \otimes_{i=1}^n \mathcal{F}_i)$ denote the product space. Then $\xi(\omega) := (\xi_1(\omega), \ldots, \xi_n(\omega))$ is a $(P, \mathcal{F}^n)$-perfect function with $P_\xi = P_{\xi_1} \otimes \cdots \otimes P_{\xi_n}$ and $\eta(\omega_1, \omega_2) := (\xi(\omega_1), \xi(\omega_2))$ is a random function from $\Omega \times \Omega$ into $\Omega^n \times \Omega^n$ with distribution $P_\eta = P_\xi \otimes P_\xi$. Let $\delta = (\delta_1, \ldots, \delta_n) \in \mathrm{E}^n$ be given. If $u = (u_1, \ldots, u_n) \in \Omega^n$ and $v = (v_1, \ldots, v_n) \in \Omega^n$ are given "vectors", we define $J_\delta(u, v) := (u^*, v^*)$ where $(u_i^*, v_i^*) = (u_i, v_i)$ if $\delta_i = 1$ and $(u_i^*, v_i^*) = (v_i, u_i)$ if $\delta_i = -1$; that is, $J_\delta : \Omega^n \times \Omega^n \to \Omega^n \times \Omega^n$ is the function which interchanges the coordinates u_i and v_i if $\delta_i = -1$ and leaves the coordinates at their original positions if $\delta_i = 1$. Hence, we see that $J_\delta : (\Omega^n \times \Omega^n, \mathcal{F}^n \otimes \mathcal{F}^n) \to (\Omega^n \times \Omega^n, \mathcal{F}^n \otimes \mathcal{F}^n)$ is a measurable bijection such that $J_\delta = J_\delta^{-1}$. So by Lem.B.6.8 we see that J_δ is $(P_\eta, \mathcal{F}^n \otimes \mathcal{F}^n)$-perfect function and since $P_\eta = P_{\xi_1} \otimes \cdots \otimes P_{\xi_n} \otimes P_{\xi_1} \otimes \cdots \otimes P_{\xi_n}$, then by the definition of J_δ we have $(P_\eta)_{J_\delta} = P_\eta$. Let $A \subseteq \Lambda^n$ be given and let us define $\tilde{X}(u) := (\hat{X}_1(u_1), \ldots, \hat{X}_n(u_n))$ for all $u = (u_1, \ldots, u_n) \in \Omega^n$ and

$$D := \{ (u, v) \in \Omega^n \times \Omega^n \mid \tilde{X}(u) - \tilde{X}(v) \in A\}$$

Since $X_i(\omega) = \hat{X}_i(\xi_i(\omega))$, we have $\{X^s \in A\} = \eta^{-1}(D)$ and since J_δ is $(P_\eta, \mathcal{F}^n \otimes \mathcal{F}^n)$-perfect with $(P_\eta)_{J_\delta} = P_\eta$, then by Lem.A.10.4 we have

$$P_2^*(X^s \in A) \leq P_\eta^*(D) = ((P_\eta)_{J_\delta})^*(D) = P_\eta^*(J_\delta^{-1}(D))$$

with equality if η is perfect; that is, if $X_1, \ldots, X_n$ are strictly perfectly independent. Since J_δ interchanges the coordinates u_i and v_i if $\delta_i = -1$ and leaves the coordinates at their original positions if $\delta_i = 1$, we have

$$J_\delta^{-1}(D) = \{ (u, v) \in \Omega^n \times \Omega^n \mid T_\delta(\tilde{X}(u) - \tilde{X}(v)) \in A \}$$

and since $T_\delta(X^s(\omega_1, \omega_2)) = T_\delta(\tilde{X}(\xi(\omega_1)) - \tilde{X}(\xi(\omega_2)))$ for all $(\omega_1, \omega_2) \in \Omega \times \Omega$, then by Lem.A.10.4 we have

$$P_2^*(X^s \in A) \leq P_\eta^*(J_\delta^{-1}(B)) = P_\eta^*((u, v) \mid T_\delta(\tilde{X}(u) - \tilde{X}(v)) \in A)$$
$$P_2^*(T_\delta(X^s) \in A) \leq P_\eta^*((u, v) \mid T_\delta(\tilde{X}(u) - \tilde{X}(v)) \in A)$$

with equality if η is perfect; that is, if $X_1, \ldots, X_n$ are strictly perfectly independent. Hence, if $X_1, \ldots, X_n$ are strictly perfectly independent, we have that X^s and $T_\delta(X^s)$ have the same outer distribution for all $\delta \in \mathrm{E}^n$ which proves (5). Let $A, B \subseteq \Lambda^n$ be given sets and let us define $F_0 := \{u \in \Omega^n \mid T_\delta(\tilde{X}(u)) \notin B\}$ and $F_1 := \{u \in \Omega^n \mid T_\delta(\tilde{X}(u)) \in A + B\}$. Let $(u, v) \in J_\delta^{-1}(D)$ be given . If $\tilde{X}(v) \in B$, then we have

$$T_\delta(\tilde{X}(u)) = T_\delta(\tilde{X}(u) - \tilde{X}(v)) + T_\delta(\tilde{X}(v)) \in A + B$$

Hence, we see that $J_\delta^{-1}(D) \subseteq (F_1 \times \Omega^n) \cup (\Omega^n \times F_0)$ and since $P_\eta = P_\xi \otimes P_\xi$ and ξ is $(P, \mathcal{F}^n)$-perfect, then by Thm.A.8 we have

$$\begin{aligned} P_2^*(X^s \in A) &\leq P_\eta^*(J_\delta^{-1}(D)) \leq P_\eta^*(F_1 \times \Omega^n) + P_\eta^*(\Omega^n \times F_0) \\ &= P_\xi^*(F_1) + P_\xi^*(F_0) = P^*(\xi^{-1}(F_1)) + P^*(\xi^{-1}(F_0)) \\ &= P^*(T_\delta(X) \notin B) + P^*(T_\delta(X) \in A + B) \end{aligned}$$

for all $\delta \in \mathrm{E}^n$ which proves (6).

(7): Evidently, we have that $X : \Omega \to (\Lambda^n, \mathcal{L}^n)$ is P-measurable and since $(\Lambda^n, \mathcal{L}^n)$ is a measurable linear space and $X^s(\omega_1, \omega_2) = X(\omega_1) - X(\omega_2)$, we see that $X^s : \Omega_2 \to (\Lambda^n, \mathcal{L}^n)$ is P_2-measurable. Let $A \in \mathcal{L}^n$, $B \subseteq \Lambda^n$ and $\delta \in \mathrm{E}^n$ be given. Since $(\Lambda, \mathcal{L})$ is a measurable linear space, we have $T_\delta^{-1}(A) \in \mathcal{L}^n$ and since X^s is $\mathcal{L}$-symmetric, we have that $P_2(X^s \in A) = P_2(T_\delta(X^s) \in A)$. Since $(\Lambda^n, \mathcal{L}^n)$

is a measurable linear space, we have that $A_2 := \{(x,y) \in \Lambda^n \times \Lambda^n \mid x - y \in A\}$ belongs to $\mathcal{L}^n \otimes \mathcal{L}^n$. Since

$$\{\,(\omega_1,\omega_1) \mid T_\delta(X^s(\omega_1,\omega_2)) \in A\} = \{\,(v,\omega) \mid T_\delta(X(\omega_1) \in A + X(\omega_2)\}$$

then by Fubini's theorem

$$P_2(X^s \in A) = P_2(T_\delta(X^s) \in A) = \int_\Omega P(T_\delta(X) \in A + T_\delta(X(\omega)))\, P(d\omega)$$

Let us define $f_\delta(\omega) := P(T_\delta(X) \in A + T_\delta(X(\omega)))$ for all $\omega \in \Omega$. If $T_\delta(X(\omega)) \in B$, we have $A + T_\delta(X(\omega)) \subseteq A + B$ and so we have $f_\delta(\omega) \le P_*(T_\delta(X) \in A + B)$. Hence, we have

$$f_\delta(\omega) \le 1_{\{T_\delta(X) \notin B\}}(\omega) + 1_{\{T_\delta(X) \notin B\}}(\omega)\, P_*(T_\delta(X) \in A + B) \quad \forall\, \omega \in \Omega$$

and so by Lem.A.2.2+6 we have

$$\begin{aligned} P_2(X^s \in A) &\le \int_* 1_{\{T_\delta(X) \notin B\}}\, dP + \int^* 1_{\{X \in B\}}\, P_*(T_\delta(X) \in A + B)\, P(d\omega) \\ &\le P_*(T_\delta(X) \notin B) + P_*(T_\delta(X) \in A + B) \end{aligned}$$

Thus, we see that $A \in \Sigma_P(X)$ for all $A \in \mathcal{L}^n$ and by Thm.A.8 and Lem.A.1.1 we see that (7.a) holds. Since $P_*(T_\delta(X) \in D) = 1 - P^*(T_\delta(X) \notin D)$ and ε and X are totally independent functions with $X_\varepsilon = T_\varepsilon(X)$, we see that (7.b) follows from Thm.2.11. $\square$

Lemma 3.5: *Let* Λ *be a vector space* $S_1 : \Omega \to \Lambda$ *and* $S_2 :\to \Lambda$ *be given functions and let* $\Psi \subseteq \bar{\mathbf{R}}^\Lambda$ *be a non-empty family of subadditive function. Let* $s, t \in \mathbf{R}$ *be given numbers and let us define* $Q(x) := \sup_{\psi \in \Psi} \psi(x)$ *for all* $x \in \Lambda$ *and*

$$D_t := \{\,(x,y) \in \Lambda \times \Lambda \mid Q(x-y) > t\} \ \ , \ \ \beta(s) := \inf_{\psi \in \Psi} P^*(\psi(S_2) \le s)$$

If $F \in \mathcal{F}_P$ *be a given set satisfying* $F \subseteq \{Q(S_1)^* > s + t\}$ *a.s., then we have (see the remark below)*

(1) *If* F *and* $\{Q(S_2)^* \le s\}$ *are non-negatively correlated, then we have*

(a) $P_*(Q(S_2) \le s)\, P^*(G) \le P^*(G \cap \{Q(S_1 - S_2) > t\})$

(2) *If* $F \cap \{Q(S_1) > s + t\}$ *and* $\{Q \le s\}$ *are non-negatively correlated, then we have*

(a) $P^*(Q(S_2) \le s)\, P(F) \le P^*(F \cap \{Q(S_1) > s + t\ ,\ Q(S_1 - S_2) > t)$

(3) *If* $(1_F, S_1)$ *and* S_2 *are upper* 1_{D_t}*-independent, then we have*

(a) $\beta(s)\, P(F) \le P^*(F \cap \{Q(S_1) > s + t\ ,\ Q(S_1 - S_2) > t)$

Remark: If $G, H \subseteq \Omega$ are given subsets of Ω, we say that G and H are *non-negatively correlated* if $P^*(G)\,P^*(H) \leq P^*(G \cap H)$.

Proof: (1+2): Since Q is subadditive, we have $Q(S_1) \leq Q(S_2) \dot{+} Q(S_1 - S_2)$ and so we have $Q(S_1)^* \leq Q(S_2)^* \dot{+} Q(S_1 - S_2)^*$. Since $F \in \mathcal{F}_P$ and $F \subseteq \{Q(S_1)^* > s+t\}$ a.s., then by Lem.A.5.5+10 we have

$$F \cap \{Q(S_2)^* \leq s\} \subseteq F \cap \{Q(S_1 - S_2)^* > t\} \text{ a.s.}$$

$$F \cap \{Q(S_2) \leq s\,,\, Q(S_1) > s + t\} \subseteq F \cap \{Q(S_1 - S_2) > t\}$$

$$P(F) = P(F \cap \{Q(S_1)^* > s + t\}) = P^*(F \cap \{Q(S_1) > s + t\})$$

$$P(F \cap \{Q(S_1 - S_2)^* > t\}) = P^*(F \cap \{Q(S_1 - S_2) > t\})$$

Hence, if F and $\{Q(S_2)^* \leq s\}$ are non-negatively correlated, then we have

$$\begin{aligned} P_*(Q(S_2) \leq s)\,P(F) = P(Q(S_2)^* \leq s)\,P(F) &\leq P(F \cap \{Q(S_2)^* \leq s\}) \\ &\leq P(F \cap \{Q(S_1 - S_2)^* > t\}) = P^*(F \cap \{Q(S_1 - S_2) > t\}) \end{aligned}$$

and if $F \cap \{Q(S_1) > s + t\}$ and $\{Q(S_2) \leq s\}$ are non-negatively correlated

$$\begin{aligned} P^*(Q(S_2) \leq s)\,P(F) &= P^*(Q(S_2) \leq s)\,P^*(F \cap \{Q(S_1) > s + t\}) \\ &\leq P^*(F \cap \{Q(S_2) \leq s\,,\, Q(S_1) > s + t\}) \\ &\leq P^*(F \cap \{Q(S_1 - S_2) > t\}) \end{aligned}$$

which proves (1) and (2).

(3): Let us define $A := \{x \in \Lambda \mid Q(x) > s + t\}$ and $f(x) := P^*(S_2 \in D_t^x)$ for all $x \in \Lambda$ where $D_t^x := \{y \in \Lambda \mid (x, y) \in D_t\}$ is the x-section of D_t . Let $x \in A$ be given. Since $Q(x) = \sup_{\psi \in \Psi} \psi(x) > s + t$, there exists a function $\psi_x \in \Psi$ such that $\psi_x(x) > s + t$. Let $y \in \Lambda$ be a given vector satisfying $\psi_x(y) \leq s$. Since ψ_x is subadditive, we have

$$s + t < \psi_x(x) = \psi_x(y) \dot{+} \psi_x(x - y) \leq s + Q(x - y)$$

So we have $\{y \in \Lambda \mid \psi_x(y) \leq s\} \in D_t^x$ and since $\beta(s) \leq P^*(\psi_x(S_2) \leq s)$, we have $f(x) \geq \beta(s)\,1_A(x)$ for all $x \in \Lambda$. Since $(1_F, S_1)$ and S_2 are upper 1_{D_t}-independent and $F \subseteq \{Q(S_1)^* > s + t\}$ a.s., then by Thm.2.6.3 and Lem.A.5.5+10 we have

$$\begin{aligned} &P^*(F \cap \{Q(S_1 - S_2) > t\}) \geq E^*\{1_F\, f(S_1)\} \geq E^*\{\beta(s)\, 1_F\, 1_A(S_1)\} \\ &= \beta(s)\, P^*(F \cap \{S_1 \in A\}) = \beta(s)\, P^*(F \cap \{Q(S_1) > s + t\}) = \beta(s)\, P^*(F) \end{aligned}$$

which proves (3.a). $\square$

4. Stochastic inequalities Recall that the classical symmetrization, Ottaviani, Lévy, exponential and subgaussian inequalities gives various upper estimates of the tail probabilities $P(|S_n| > s)$ or $P(M_n > s)$ where S_n is the sum of the independent random variables $X_1, \ldots, X_n$ and $M_n := \max_{1\leq i\leq n} |X_1 + \cdots + X_i|$ is the maximal absolute partial sum. In this section we shall study extensions of the classical inequalities to "norms" of sums arbitrary vector-valued functions. More precisely, let Λ be a vector space and let $X_1, \ldots, X_n : \Omega \to \Lambda$ and $Q : \Lambda \to \bar{\mathbf{R}}$ be given functions. Let $S_i := X_1 + \cdots + X_i$ denote the partial sums for $i = 1, \ldots, n$ and let $M_n := \max_{1\leq i\leq n} Q(S_i)$ denote the maximal "norm" of the partial sums. Then we shall search for upper estimates of the tail probabilities $P^*(Q(S_n) > s)$ or $P^*(M_n > s)$. In the current literature (see e.g. [3], [4], [12], [13], [11] and [15]), the extensions of the classical inequalities are proved under one of the following (implicitly or explicitly stated) assumptions

(4.A) Q is a seminorm (or a convex increasing function of a seminorm) and the functions $X_1, \ldots, X_n$ are independently defined

(4.B) $(\Lambda, \mathcal{L})$ is a measurable linear space such that Q is an $\mathcal{L}$-measurable seminorm (or a convex increasing function of an $\mathcal{L}$-measurable seminorm) and $X_1, \ldots, X_n : \Omega \to (\Lambda, \mathcal{L})$ are independent random functions

and in some cases with additional measurability assumptions (e.g. the subgaussian inequality) or additional symmetry assumptions (e.g. Lévy's inequality) or under the assumption that $X_1, \ldots, X_n$ are stochastic processes.

Let T be a non-empty set. Observe that the sup-norm $||\cdot||_T$ is $\mathcal{B}^T$-measurable if and only if T is countable, which means that for an uncountable index set T, we are led to the study of non-measurable functions. A standard method in the current literature is to replace the sup-norm with an appropriate upper envelope (which will be measurable). However, if T is uncountable and μ is any given probability measure on $(\mathbf{R}^T, \mathcal{B}^T)$ and $||\cdot||_T^*$ is any upper μ-envelope of the sup-norm, then the reader easily verifies that $||f||_T^* = \infty$ μ-a.s. In the literature, this problem has been bypassed by assuming that $X_1, \ldots, X_n$ are independently defined and choosing a more appropriate (and non-trivial) envelope of the sup-norm (using perfectness of projections, see Lem.B.7). The observation shows that perfectness of a stochastic process with an uncountable parameter set is very restrictive assumption; for instance, if T is uncountable and $X : \Omega \to (\mathbf{R}^T, \mathcal{B}^T)$ is a stochastic process with a.s. bounded sample paths or more generally with $P^*(||X||_T < \infty) > 0$, then X is *not* $(P, \mathcal{B}^T)$-perfect – so assuming perfectness of the stochastic processes $X_1, \ldots, X_n$ (which makes the inequalities easy consequences of the measurable case) is far too restrictive.

In this section, we shall see that most of the classical inequalities do hold for a large class of seminorm-like functions $Q : \Lambda \to \bar{\mathbf{R}}$ and under much weaker independence properties of the functions $X_1, \ldots, X_n$ than (4.A) or (4.B). The following definition clarifies the term "seminorm-like function":

Definition 4.1: (Convexity) Let Λ be a vector space and let $\Xi \subseteq \Lambda^*$ be a given set of linear functionals. Then we let $\mathrm{span}(\Xi)$ denote *the linear span* of Ξ in the dual space Λ^* and we let $\mathrm{aff}(\Xi)$ denote *the affine span* of Ξ ; that is, the set of all functions of the form $x \curvearrowright \xi(x) + c$ for some $\xi \in \mathrm{span}(\Xi)$ and some $c \in \mathbf{R}$. We let $\Xi^\perp := \{x \in \Lambda \mid \xi(x) = 0 \ \ \forall \xi \in \Xi\}$ denote *the orthogonal complement* of Ξ . If $\psi : \Lambda \to \bar{\mathbf{R}}$ is a given function and $\Xi \subseteq \Lambda^*$ is a non-empty set of linear functions, we say that ψ is

- *convex* if $\psi(\alpha x + (1-\alpha)\, y) \leq \alpha\, \psi(x) \dot{+} (1-\alpha)\, \psi(y) \ \ \forall\, x, y \in \Lambda\ \forall\, 0 \leq \alpha \leq 1$
- *quasiconvex* if $\psi(\frac{1}{2}\, x + \frac{1}{2}\, y) \leq \psi(x) \vee \psi(y) \ \ \forall\, x, y \in \Lambda$
- Ξ-*convex* if $\exists\, \Phi \subseteq \mathrm{aff}\,(\Xi)$ so that $\Phi \neq \emptyset$ and $\psi(x) = \sup_{\phi \in \Phi} \phi(x) \ \ \forall\, x \in \Lambda$
- *even* if $\psi(x) = \psi(-x) \ \ \forall\, x \in \Lambda$

Let $[\alpha, \beta] \subseteq \bar{\mathbf{R}}$ be a closed (possibly unbounded) interval and let $f : [\alpha, \beta] \to \bar{\mathbf{R}}$ be an increasing continuous function such that f is finite and convex on the open interval (α, β) . Then we have the following well-known (and simple) facts:

(1) If $\phi : \Lambda \to \bar{\mathbf{R}}$ is Ξ-convex and $C \subseteq \Lambda$ is a non-empty set such that $\phi(x) \in [a, b]$ for all $x \in C$, then there exists a Ξ-convex function $\psi : \Lambda \to \bar{\mathbf{R}}$ such that $\psi(x) = f(\phi(x))$ for all $x \in C$

(2) The sets of all subadditive, all convex, all Ξ-convex, all quasiconvex and all even functions on Λ are stable under arbitrary suprema

(3) The sets of all convex and all Ξ-convex functions on Λ are convex cones containing all constant functions

(4) If $\phi : \Lambda \to \bar{\mathbf{R}}$ is Ξ-convex, then ϕ is convex and we have

   (a) $-\infty < \phi(x) \leq \infty \quad \forall\, x \in \Lambda$

   (b) $\phi(x) = \phi(y) \quad \forall\, x, y \in \Lambda$ with $x - y \in \Xi^\perp$

(5) If Λ is a locally convex linear topological space and $\phi : \Lambda \to (-\infty, \infty]$ is a lower semicontinuous, convex function, then ϕ is Λ'-convex where Λ' denotes the *topological dual space*; that is, the set of all continuous, real-valued, linear functionals on Λ

Let $Q : \Lambda \to [0, \infty]$ be given seminorm on Λ . Then Q is a non-negative, convex, subadditive and even function on Λ and if $B := \{x \in \Lambda \mid Q(x) \leq 1\}$ denotes *the Q-unit ball* and $B^\circ := \{\xi \in \Lambda^* \mid |\xi(x)| \leq 1\ \forall\, x \in B\}$ denotes *the polar* of B , then by Hahn-Banach's theorem we have

(6) $Q(x) = \sup\limits_{\xi \in B^\circ} \xi(x) = \sup\limits_{\xi \in B^\circ} |\xi(x)| \quad \forall\, x \in \Lambda$

In particular, we see that Q is B°-convex.

Let T be a non-empty set and let us consider the measurable linear space $(\mathbf{R}^T, \mathcal{B}^T)$ equipped with its product topology. Then $\mathbf{R}^T$ is a locally convex K-analytic space (see [14]) and $\mathcal{B}^T$ is the Baire σ-algebra on $\mathbf{R}^T$ (*the Bockstein–Kellerer theorem*; see [2] and [10]). We let $\pi_t(f) := f(t)$ for $f \in \mathbf{R}^T$ and $t \in T$ denote the *projections maps*. Note that π_t is a continuous, $\mathcal{B}^T$-measurable, linear function on $\mathbf{R}^T$ for all $t \in T$ and if we define $\Pi := \{\pi_t \mid t \in T\}$ and $\Pi^* := \Pi \cup \{-\pi_t \mid t \in T\}$, then we have

(7) $\quad \sup_T(f) = \sup_{\pi\in\Pi} \pi(f) \quad , \quad \|f\|_T = \sup_{\pi\in\Pi} |\pi(f)| = \sup_{\pi\in\Pi^*} \pi(f) \quad \forall\, f \in \mathbf{R}^T$

In particular, we see that the sup-functional $\sup_T$ is lower semicontinuous, Π-convex and subadditive and that the sup-norm $||\cdot||_T$ is a lower semicontinuous, Π^*-convex seminorm.

Definition 4.2: (The mean of random vector) Let Λ be a vector space and let $\Xi \subseteq \Lambda^*$ be a given set of linear functions. If $X : \Omega \to \Lambda$ is a given function, we say that X is Ξ-*integrable* if and only if

(1) $\quad \xi(X) \in L^1(P) \ \ \forall\, \xi \in \Xi$ and $\exists\, a \in \Lambda$ so that $\xi(a) = E\xi(X) \ \ \forall\, \xi \in \Xi$

The set of all $a \in \Lambda$ satisfying $\xi(a) = E\xi(X) \ \forall\, \xi \in \Xi$ will be denoted $\mathcal{E}_\Xi(X)$ and any vector $a \in \mathcal{E}_\Xi(X)$ will be called an Ξ-*mean* of X . If X is Ξ-integrable and $0 \in \mathcal{E}_\Xi(X)$, we say that X has *zero* Ξ-*mean*. Note that X has Ξ-mean zero if and only if $\xi(X) \in L^1(P)$ and $E\xi(X) = 0$ for all $\xi \in \Xi$. If $X : \Omega \to \Lambda$ is a Ξ-integrable function, then the reader easily verifies the following simple facts:

(2) $\quad \psi(X) \in L^1(P)$ and $\psi(a) = E\psi(X) \ \ \forall\, \psi \in \operatorname{aff}(\Xi) \ \forall\, a \in \mathcal{E}_\Xi(X)$

(3) $\quad \Xi^\perp = \{\, a - b \mid a, b \in \mathcal{E}_\Xi(X)\}$

(4) $\quad \mathcal{E}_\Xi(X) = \mathcal{E}_{\operatorname{span}(\Xi)}(X) = \{\, a + x \mid x \in \Xi^\perp\} \ \ \forall\, a \in \mathcal{E}_\Xi(X)$

Hence, if $\Xi^\perp = \{0\}$, then by (3+4) we see that the Ξ-mean of X is unique but if $\Xi^\perp \neq \{0\}$, then the set of Ξ-means of X is a non-trivial affine subspace of Λ .

Let T be a non-empty set and let us consider the vector space $\mathbf{R}^T$. If $\Pi := \{\pi_t \mid t \in T\}$ denotes the set of all projections, then a stochastic process $X = (X_t)$ is Π-integrable if and only if $X_t \in L^1(P)$ for all $t \in T$ and if so, then we have $\mathcal{E}_\Pi(X) = \{\mu\}$ where $\mu(t) := EX_t$ is *the mean value function*.

Theorem 4.3: (Jensen's inequality) *Let Λ be a vector space and let $\Xi \subseteq \Lambda^*$ be a non-empty set of linear functions. Let $\phi : \Lambda \to \bar{\mathbf{R}}$ be a Ξ-convex function and let $X : \Omega \to \Lambda$ be a Ξ-integrable function. Then we have*

(1) $\quad \phi(a) = \phi(b) \leq E_*\phi(X) \quad \forall\, a, b \in \mathcal{E}_\Xi(X)$

Let $Y:\Omega\to\Lambda$ *be a given function and let us define* $\psi(x,y):=\phi(x+y)$ *for all* $(x,y)\in\Lambda\times\Lambda$ *. If* Y *and* X *are upper* ψ*-independent, then we have*

(2) $\quad E^*\phi(Y+a)\} \le E^*\phi(X+Y) \quad \forall\, a\in\mathcal{E}_\Xi(X)$

Remarks: (a): Suppose that Y and X are totally independent. By Thm.2.6.1 we have that Y and X are upper h-independent for all $h\in\bar{\mathbf{R}}^\Lambda$. Hence, we see that (2) holds.

(b): Let $(\Lambda,\mathcal{L})$ be a measurable linear space such that $X,Y:\Omega\to(\Lambda,\mathcal{L})$ are independent random functions and $\psi(x,y):=\phi(x+y)$ is $P_{Y,X}$-measurable. By Def.2.5.2 we have that Y and X are upper ψ-independent. Hence, we see that (2) holds.

Proof: (1): Since ϕ is Ξ-convex, there exists a non-empty set $\Psi\subseteq\operatorname{aff}\Xi$ such that $\phi=\sup_{\psi\in\Psi}\psi$ and since $a,b\in\mathcal{E}_\Xi(X)$, then by Def.4.2.2+4 we have $\phi(a)=\phi(b)$. Let $\psi\in\Psi$ be given. Since $\psi(u)\le\phi(u)$ and $\psi(X)\in L^1(P)$, then by Def.4.2.2 we have $\psi(a)=E\psi(X)\le E_*\phi(X)$. Taking supremum over $\psi\in\Psi$, we obtain (1).

(2): Let $x\in\Lambda$ be given. Since $x+a\in\mathcal{E}_\Xi(x+X)$, then by (1) we have $\phi(x+a)\le E_*\phi(x+X)\le E^*\phi(x+X)$ and since Y and X are upper ψ-independent, we see that (2) holds. $\square$

Theorem 4.4: (The symmetrization inequalities) *Let* Λ *be a vector space, let* $X=(X_1,\dots,X_n)$ *be a function from* Ω *into* Λ^n *and let us define* $S_\delta:=\delta_1X_1+\cdots+\delta_nX_n$ *for all* $\delta=(\delta_1,\dots,\delta_n)\in\mathrm{E}^n$ *. Let* $\Psi\subseteq\bar{\mathbf{R}}^\Lambda$ *be a non-empty family subadditive functions and let us define* $Q(x):=\sup_{\psi\in\Psi}\psi(x)$ *for all* $x\in\Lambda$ *. Let* $\varepsilon=(\varepsilon_1,\dots,\varepsilon_n)$ *be a random function from* Ω *into* $(\mathrm{E}^n,\mathcal{E}^n)$ *such that*

(1) $\quad \varepsilon$ *and* $Q(S_\delta)^*$ *are independent and* ε *and* $Q(-S_\delta)^*$ *are independent* $\forall\,\delta\in\mathrm{E}^n$

Let $X^s=(X_1^s,\dots,X_n^s)$ *and* $X^\epsilon=(X_1^\epsilon,\dots,X_\epsilon^s)$ *denote the standard and Bernoulli symmetrizations of* X *and let us define*

$$S:=\sum_{i=1}^n X_i\ ,\quad S^s:=\sum_{i=1}^n X_i^s\ ,\quad S^\epsilon:=\sum_{i=1}^n X_i^\epsilon\ ,\quad S_\varepsilon:=\sum_{i=1}^n\varepsilon_iX_i\ ,\quad Q(x):=\sup_{\psi\in\Psi}\psi(x)$$

$$(\Omega_2,\mathcal{F}_2,P_2):=(\Omega\times\Omega,\mathcal{F}\otimes\mathcal{F},P\otimes P)\ ,\quad (\tilde\Omega,\tilde{\mathcal{F}},\tilde P):=(\mathrm{E}^n\times\Omega,\mathcal{E}^n\otimes\mathcal{F},\mathrm{B}^n\otimes P)$$

$$\Delta_s:=\{\,(x_1,\dots,x_n)\in\Lambda^n\mid Q(x_1+\cdots+x_n)>s\}\ ,\quad \beta(s):=\inf_{\psi\in\Psi}P^*(\psi(S)\le s)$$

for all $x\in\Lambda$ *and all* $s\in\mathbf{R}$ *. If* $F\in\mathcal{C}_P(X)$ *is any given set, then we have*

(2) $\quad \beta(s)P^*(F\cap\{Q(S)>s+t\})\le P_2^*((F\times\Omega)\cap\{Q(S^s)>t\})\ \ \forall s,t\in\mathbf{R}$

(3) $\quad P_2^*(Q(S^s)>t+u)\le P^*(Q(-S)>t)+P^*(Q(S)>u)\quad\forall\,t,u\in\mathbf{R}$

(4) *If* $t, u \in \mathbf{R}$ *are given numbers satisfying* $\Delta_{t+u} \in \Sigma_P(X)$ *, then we have*

(a) $P_2^*(Q(S^s) > t+u) \le \tilde{P}^*(Q(S^\epsilon) > t) + \tilde{P}^*(Q(S^\epsilon) > u)$

(b) $P_2^*(Q(S^s) > t+u) \le P^*(Q(-S_\varepsilon) > t) + P^*(Q(S_\varepsilon) > u)$

(c) $\beta(s)\, P^*(Q(S) > s+t+u) \le \tilde{P}^*(Q(S^\epsilon) > t) + \tilde{P}^*(Q(S^\epsilon) > u) \ \ \forall\, s \in \mathbf{R}$

(d) $\beta(s)\, P^*(Q(S) > s+t+u) \le P^*(Q(-S_\varepsilon) > t) + P^*(Q(S_\varepsilon) > u) \ \ \forall\, s \in \mathbf{R}$

Remarks: (a): If ε and X are totally independent, then by Lem.2.1.5 and Thm.2.9.1 we see that condition (1) holds. In particular, we see that $\varepsilon \equiv \delta$ satisfies condition (1) for all $\delta \in \mathrm{E}^n$. If $Q(S_\delta)$ and $Q(-S_\delta)$ are P-measurable and independent of ε for all $\delta \in \mathrm{E}^n$, then condition (1) holds trivially.

(b): Observe that the symmetrization inequalities (1), (2) and (3.a)–(3.d) hold far beyond the standard hypotheses (4.A) or (4.B) and recall that Thm.3.4.4–7 provide several criteria implying $\Delta_{t+u} \in \Sigma_P(X)$. In particular, we see that the symmetrization the inequalities (4.a)–(3.d) hold for all $s, t, u \in \mathbf{R}$ in each of the following three cases. (i): X^s is totally symmetric; (ii): $X_1, \ldots, X_n$ are perfectly independent; (iii): $(\Lambda, \mathcal{L})$ is a measurable linear space such that Q is $\mathcal{L}$-measurable and $X_1, \ldots, X_n : \Omega \to (\Lambda, \mathcal{L})$ are independent random functions.

(c): Let $X_1, \ldots, X_n : \Omega \to (\mathbf{R}^T, \mathcal{B}^T)$ be stochastic processes and let $S := X_1 + \cdots + X_n$ denote the sum process. By Def.4.1.7 we have following expressions for the function $\beta(s)$ when Q is the sup-functional or the sup-norm

(5) $Q(f) = \sup_T f \quad \Rightarrow \quad \beta(s) = \inf_{t \in T} P(S(t) \le s) \ \ \forall\, s \in \mathbf{R}$

(6) $Q(f) = \|f\|_T \quad \Rightarrow \quad \beta(s) = \inf_{t \in T} \{P(S(t) \le s) \wedge P(S(t) \ge -s)\} \ \ \forall\, s \in \mathbf{R}$

Proof: (2+3): Let $s, t \in \mathbf{R}$ be given and let us define $Y_i(\omega_1, \omega_2) := X(\omega_i)$ and $S_i(\omega_1, \omega_2) := S(\omega_i)$ for $i = 1, 2$ and $(\omega_1, \omega_2) \in \Omega \times \Omega$. Since $Y_1^{-1}(A) = X^{-1}(A) \times \Omega$ for all $A \subseteq \Lambda^n$, we see that $H \times \Omega \in \mathcal{S}_{P_2}(Y_1)$ for all $H \in \mathcal{S}_P(X)$ and by Thm.A.8 we have $(H \times \Omega)_* = H_* \times \Omega$ for all $H \subseteq \Omega$ where H_* denotes the P-kernel of H and $(H \times \Omega)_*$ denotes the P_2-kernel of $H \times \Omega$. Hence, we see that $H \times \Omega \in \mathcal{C}_{P_2}(Y_1)$ for all $H \in \mathcal{C}_P(X)$ and since $F \in \mathcal{C}_P(X)$, we have that $F \times \Omega \in \mathcal{C}_{P_2}(Y_1)$. Since Y_1 and Y_2 are independently defined, then by Thm.2.10.4 and Thm.2.6.2 we have that $(1_{F \times \Omega}, S_1)$ and S_2 are totally independent and by Thm.A.8 we have

$$\beta(s) = \inf_{\psi \in \Psi} P^*(\psi(S) \le s) = \inf_{\psi \in \Psi} P_2^*(\psi(S_2) \le s) \quad \forall\, s \in \mathbf{R}$$

Since $S^s = S_1 - S_2$, we see that (2) follows from Lem.3.5.3 and Thm.A.8. (3) is an immediate consequence of Thm.A.8 and subadditivity of Q .

(4): Suppose that $\Delta_{t+u} \in \Sigma_P(X)$ and let us define $B_u := \{x \in \Lambda^n \mid Q_n(-x) \le u\}$ and $Q_n(x) := Q(x_1 + \cdots + x_n)$ for all $x = (x_1, \ldots, x_n) \in \Lambda^n$. Let $x \in \Delta_{t+u} + B_u$ be given. Then the exists $y \in \Delta_{s+t}$ and $z \in B_u$ such that $x = y + z$ and since Q_n is subadditive, we have

$$t + u \le Q_n(y) = Q_n(x - z) \le Q_n(x) \dot{+} Q_n(-z) \le Q_n(x) + u$$

Hence, we see that $\Delta_{t+u} + B_u \subseteq \Delta_t$. Since X^ϵ is totally symmetric (see Thm.3.4.3), we see that (4.a) follows from Thm.3.4.1. Let $\delta \in \mathrm{E}^n$ be given. Since a constant function is totally independent of any function, then by Thm.3.4.2 we have

$$P_2^*(Q(S^s) > t+u) \le P^*(Q(-S_\delta) > t) + P^*(Q(S_\delta) > u) \quad \forall\, \delta \in \mathrm{E}^n$$

Hence, by we see that (4.b) follows from (1) and Thm.2.11.1. The inequalities (4.c+d) are immediate consequences of (2), (4.a) and (4.b). □

Theorem 4.5: (The convex symmetrization inequalities) *Let Λ be a vector space, let $X = (X_1, \dots, X_n)$ be function from Ω into Λ^n and let us define $S_\delta := \delta_1 X_1 + \cdots + \delta_n X_n$ for all $\delta = (\delta_1, \dots, \delta_n) \in \mathrm{E}^n$. Let $\Xi \subseteq \Lambda^*$ be a non-empty set of linear functional such that $X_1, \dots, X_n$ have zero Ξ-means. Let $Q : \Lambda \to \bar{\mathbf{R}}$ be an even Ξ-convex function and let $\varepsilon = (\varepsilon_1, \dots, \varepsilon_n)$ be a random function from Ω into $(\mathrm{E}^n, \mathcal{E}^n)$ satisfying*

(1) *ε and $Q(S_\delta + \mu)^*$ are independent for all $\delta \in \mathrm{E}^n$ and all $\mu \in \Lambda$*

Let us define $Q_2(x,y) := Q(x+y)$ for all $(x,y) \in \Lambda^2$ and $S := \sum_{i=1}^n X_i$ and $S^\alpha := \sum_{i\in\alpha} X_i$ for all $\alpha \subseteq \{1, \dots, n\}$ with the convention $S^\emptyset := 0$. Then we have

(2) *If $2S^\alpha$ and $2\,(S - S^\alpha)$ are upper Q_2-independent for all $\alpha \subseteq \{1, \dots, n\}$, then we have*

  (a) $\quad E^*Q\Big(\sum\limits_{i=1}^n \varepsilon_i\, X_i\Big) \le E^*Q\Big(2 \sum\limits_{i=1}^n X_i\Big)$

(3) *If $2\,(S_\alpha - \mu)$ and $2\,(S - S^\alpha)$ are upper Q_2-independent for all $\alpha \subseteq \{1, \dots, n\}$ and all $\mu \in \Lambda$, then we have*

  (a) $\quad E^*Q\Big(\sum\limits_{i=1}^n X_i\Big) \le E^*Q\Big(2 \sum\limits_{i=1}^n \varepsilon_i\,(X_i - \mu_i)\Big) \quad \forall\, \mu_1, \dots, \mu_n \in \Lambda$

(4) *Let $X^s = (X_1^s, \dots, X_n^s)$ denote the standard symmetrization of X and let E_2^* denote the upper $(P \otimes P)$-expectation. If X^s satisfies the following symmetry condition*

  (a) $\quad E_2^*Q\Big(\sum\limits_{i=1}^n X_i^s\Big) = E_2^*Q\Big(\sum\limits_{i=1}^n \delta_i\, X_i^s\Big) \quad \forall\, (\delta_1, \dots, \delta_n) \in \mathrm{E}^n$

  then the inequalities (2.a) and (3.a) hold

Remarks: (a): If ε and X are totally independent, then by Lem.2.1.5 and Thm.2.9.1 we see that condition (1) holds. In particular, we see that $\varepsilon \equiv \delta$ satisfies condition (1) for all $\delta \in \mathrm{E}^n$. If $Q(S_\delta - \mu)$ is P-measurable and independent of ε for all $(\delta, \mu) \in \mathrm{E}^n \times \Lambda$, then condition (1) holds trivially.

(b): If $X_1, \dots, X_n$ have zero Ξ-means and ε satisfies condition (1), then by Thm.2.10.4 and Thm.3.4.5 we see that the inequalities (2.a) and (3.a) hold in each of the following four cases. (i): $X_1, \dots, X_n$ are totally independent; (ii): X^s is totally symmetric; (iii): $X_1, \dots, X_n$ are strictly perfectly independent; (iv): $(\Lambda, \mathcal{L})$ is a measurable linear space such that Q is $\mathcal{L}$-measurable and $X_1, \dots, X_n : \Omega \to (\Lambda, \mathcal{L})$ are independent random functions. In particular, we see that the inequalities (2.a) and (3.a) far beyond the standard hypotheses (4.A) or (4.B).

(c): Let T be a non-empty set and let $\pi_t(f) := f(t)$ denote the projections on $\mathbf{R}^T$. Let $\phi : [0, \infty] \to \mathbf{R}$ be an continuous function such that ϕ is convex and finite on $[0, \infty)$. Then $Q(f) := \phi(||f||_T)$ is a Ξ-convex function where $\Xi := \{\pm \pi_t \mid t \in T\}$ and if $X = (X_t)$ is a stochastic process, then $X : \Omega \to (\mathbf{R}^T, \mathcal{B}^T)$ is a random function and X has zero Ξ-mean if and only if $X_t \in L^1(P)$ and $EX_t = 0$ for all $t \in T$. Hence, we see that the inequalities (2.a) and (3.a) applies to stochastic processes under the appropriate independence or symmetry hypotheses.

Proof: Since Q is an even Ξ-convex function, we have $-\infty < Q(0) = Q(\frac{1}{2}(u-u)) \leq \frac{1}{2}\, Q(u) + \frac{1}{2}\, Q(-u) = Q(u)$ for all $u \in \Lambda$. Hence, we see that $q := Q(0)^-$ is a finite, non-negative number such that $Q(u) + q \geq 0$ for all $u \in \Lambda$. Thus, without loss of generality, we may assume that $Q \geq 0$.

(2): Let $\delta = (\delta_1, \dots, \delta_n) \in \mathrm{E}^n$ be a given vector and let us define $\alpha := \{i \mid \delta_i = 1\}$ and $\beta := \{i \mid \delta_i = -1\}$. Then we have $S = S^\alpha + S^\beta$ and $S_\delta = S^\alpha - S^\beta$. Since X_i has zero Ξ-mean for all $i = 1, \dots, n$, we see that $2S^\alpha$ and $2S^\beta$ have zero Ξ-means. Since $2S^\beta = 2(S - S^\alpha)$ and $2S^\alpha = 2(S - S^\beta)$, then by the assumption we see that $2S^\alpha$ and $2S^\beta$ are upper Q_2-independent and that $2S^\beta$ and $2S^\alpha$ are upper Q_2-independent. Since Q is Ξ-convex, then by Thm.4.3.2 we have $E^*Q(2S^\alpha) \leq E^*Q(2S)$ and $E^*Q(2S^\alpha) \leq E^*Q(2S)$ and since $2S_\delta = 2(S^\alpha - S^\beta)$ and Q is convex and even, then by Lem.A.5.1 we have

$$
\begin{aligned}
E^*Q(S_\delta) &\leq E^*\big\{ \tfrac{1}{2}\, Q(2S^\alpha) + \tfrac{1}{2}\, Q(2S^\beta) \big\} \\
&\leq \tfrac{1}{2}\, E^*Q(2S^\alpha) + \tfrac{1}{2}\, E^*Q(2S^\beta) \big\} \leq E^*Q(2S)
\end{aligned}
$$

for all $\delta \in \mathrm{E}^n$. Since $Q \geq 0$, we see that (2.a) follows from (1) and Thm.2.11.1.

(3): Let $\delta = (\delta_1, \dots, \delta_n) \in \mathrm{E}^n$ be given and let α and β be defined as in the proof of (2). Let $\mu_1, \dots, \mu_n \in \Lambda$ be given and let us define $\mu_\delta := \delta_1\, \mu_1 + \cdots + \delta_n\, \mu_n$. By convexity of Q , we have

$$
E^*Q(S) \leq \tfrac{1}{2}\, E^*Q(2(S^\alpha - \mu_\delta)) + \tfrac{1}{2}\, E^*Q(2(S^\beta + \mu_\delta))
$$

Since $S^\alpha = S - S^\beta$, then by assumption we have that $2(S^\beta + \mu_\delta)$ and $2S^\alpha$ are upper Q_2-independent and since $0 \in \mathcal{E}_\Xi(-2S^\alpha)$ and $2(\mu_\delta - S_\delta) = 2(S^\beta + \mu_\delta) - 2S^\alpha$, then by Thm.4.3.2 and Ξ-convexity of Q we have $E^*Q(2(S^\beta + \mu_\delta)) \leq E^*Q(2(\mu_\delta - S_\delta))$. Since $S^\beta = S - S^\alpha$, then by assumption we have that $2(S^\alpha - \mu_\delta)$ and $2S^\beta$ are upper Q_2-independent and since $0 \in \mathcal{E}_\Xi(-2S^\beta)$ and $2(S_\delta - \mu_\delta) = 2(S^\alpha + \mu_\delta) - 2S^\beta$, then by Thm.4.3.2 and Ξ-convexity of Q we have $E^*Q(2(S^\alpha - \mu_\delta)) \leq E^*Q(2(S_\delta - \mu_\delta))$. Since Q is an even function, we have

$$
E^*Q(S) \leq \tfrac{1}{2}\, E^*Q(2(S^\alpha - \mu_\delta)) + \tfrac{1}{2}\, E^*Q(2(S^\beta + \mu_\delta)) \leq E^*Q(2(S_\delta - \mu_\delta))
$$

for all $\delta \in \mathrm{E}^n$. Since $Q \geq 0$, we see that (3.a) follows from (1) andThm.2.11.1.

(4): Let $\delta \in \mathrm{E}^n$ be given and let us define $S^i(\omega_1,\omega_2) = S(\omega_i)$ and $S^i_\delta(\omega_1,\omega_2) := S_\delta(\omega_i)$ for all $(\omega_1,\omega_2) \in \Omega \times \Omega$ and $i = 1,2$. Then S^1_δ and S^2_δ are independently defined and by Thm.A.8, we have that S^1_δ and S^2_δ have zero Ξ-means. Suppose that (4.a) holds. Since Q is even and convex, then by Thm.A.8 and Thm.4.3.2 we have

$$\begin{aligned} E^*Q(S_\delta) &= E^*_2 Q(S'_\delta) \leq E^*_2 Q(S'_\delta - S''_\delta) = E^*_2 Q(S' - S'') \\ &\leq \tfrac{1}{2}\, E^*_2 Q(2S') + \tfrac{1}{2}\, E^*_2 Q(2S'') = E^*Q(2S) \end{aligned}$$

Hence, we see that (2.a) follows from (1) and Thm.2.11.1. So suppose that (4.b) holds and let $\mu_1,\dots,\mu_n \in \Lambda$ be given. Let $\delta = (\delta_1,\dots,\delta_n) \in \mathrm{E}^n$ be given and let us define $\mu_\delta := \delta_1\mu_1 + \cdots + \delta_n\mu_n$. Since Q is even and convex, then by (4.b), Thm.A.8 and Thm.4.3.2 we have

$$\begin{aligned} E^*Q(S) &= E^*_2 Q(S') \leq E^*_2 Q(S' - S'') = E^*_2 Q\big((S'_\delta - \mu_\delta) - (S''_\delta - \mu_\delta)\big) \\ &\leq \tfrac{1}{2}\, E^*_2 Q(2(S'_\delta - \mu_\delta)) + \tfrac{1}{2}\, E^*_2 Q(2(S''_\delta - \mu_\delta)) = E^*Q(2(S_\delta - \mu_\delta)) \end{aligned}$$

Hence, we see that (3.a) follows from (1) and Thm.2.11. □

Theorem 4.6: (Ottavani's inequality) *Let Λ be a vector space and let $Q : \Lambda \to \bar{\mathbf{R}}$ and $X_1,\dots,X_n : \Omega \to \Lambda$ be given functions where $n \geq 2$ is a given integer. Let $S_i := X_1 + \cdots + X_i$ denote the partial sums for $1 \leq i \leq n$ and let Q_i be a given upper P envelope of $Q(S_i)$ for all $1 \leq i \leq n$. Let us define*

$$\tau_s(\omega) = \min\{\, 1 \leq i \leq n \mid Q_i(\omega) > s \} \quad \forall\, \omega \in \Omega \quad (\ \min \emptyset := \infty\ )$$

$$D_t = \{(x,y) \in \Lambda^2 \mid Q(x-y) > t\} \ \ , \ \ \beta(s,t) = \min_{1 \leq i < n} P(Q_n > t \mid \tau_{s+t} = i)$$

for all $s,t \in \mathbf{R}$ with the convention $P(F \mid G) := 1$ if $F,G \in \mathcal{F}_P$ and $P(G) = 0$. Let $s \geq 0$ and $t \in \mathbf{R}$ be given numbers. Then we have the following Ottaviani inequalities:

(1) $\beta(s,t)\, P^*\Big(\max_{1 \leq i \leq n} Q(S_i) > s + t \Big) \leq P^*(Q(S_n) > t)$

(2) *If Q is subadditive and $\{\tau_{s+t} = i\}$ and $\{Q(S_i - S_n)^* \leq s\}$ are non-negatively correlated for all $1 \leq i < n$, then we have*

  (a) $\min_{1 \leq i < n} P_*(Q(S_i - S_n) \leq s) \leq \beta(s,t)$

(3) *If Q is subadditive and $\{\tau_{s+t} = i\ ,\ Q(S_i - S_n) \leq s\}$ and $\{Q(S_i - S_n) \leq s\}$ are non-negatively correlated for all $1 \leq i < n$, then we have*

  (a) $\min_{1 \leq i < n} P^*(Q(S_i - S_n) \leq s) \leq \beta(s,t)$

(4) *If* $Q = \sup_{\psi\in\Psi} \psi$ *for some non-empty set* $\Psi \subseteq \bar{\mathbf{R}}^\Lambda$ *of subadditive function and* $(1_{\{\tau_{s+t}=i\}}, S_i)$ *and* $S_i - S_n$ *are upper* 1_{D_t}*-independent for all* $1 \le i < n$ *, then we have*

(a) $\min_{1\le i<n} \inf_{\psi\in\Psi} P^*(\psi(S_i - S_n) \le s) \le \beta(s,t)$

Remark: (a): Note that $\{\tau_u = i\} = \{Q_i > u \, , \, Q_\nu \le u \; \forall\, 1 \le \nu < i\}$ for all $u \in \mathbf{R}$ and all $i = 1, \ldots, n$ and since Q_ν is an upper P-envelope of $Q(X_1 + \cdots + X_\nu)$, then by Lem.2.1.5 we have

(5) $\{\tau_u = i\} \in \mathcal{A}_P(X_1, \ldots, X_i) \cap \mathcal{L}_P(X_1, \ldots, X_i) \cap \sigma(Q_1, \ldots, Q_i)$

(6) $\{\tau_u = i \, , \, Q(S_i) \in A\} \in \mathcal{A}_P(X_1, \ldots, X_i) \quad \forall\, A \subseteq \bar{\mathbf{R}}$

(7) $\{Q(S_i - S_n) \in A\} \in \mathcal{S}_P(X_{i+1}, \ldots, X_n) \subseteq \mathcal{A}_P(X_{i+1}, \ldots, X_n) \quad \forall\, A \subseteq \bar{\mathbf{R}}$

for all $u \in \mathbf{R}$ and all $i = 1, \ldots, n$. Let $\Psi \subseteq \bar{\mathbf{R}}^\Lambda$ be a given family of subadditive functions and let β_* , β^* and β_Ψ denote the functions on the left hand sides of (2)–(4):

(8) $\beta_*(s) = \min_{1\le i<n} P_*(Q(S_i - S_n) \le s) \;\; , \;\; \beta^*(s) = \min_{1\le i<n} P^*(Q(S_i - S_n) \le s)$

(9) $\beta_\Psi(s) = \min_{1\le i<n} \inf_{\psi\in\Psi} P^*(\psi(S_i - S_n) \le s)$

for all $s \in \mathbf{R}$. Then we have $\beta_*(s) \le \beta^*(s)$ for all $s \in \mathbf{R}$ and if $Q = \sup_{\psi\in\Psi} \psi$, then we have $\beta^*(s) \le \beta_\Psi(s)$ for all $s \in \mathbf{R}$. Moreover, under the appropriate independence conditions we have that Ottaviani's inequality (1) holds with $\beta(s,t)$ replaced by $\beta_*(s)$ or $\beta^*(s)$ or $\beta_\Psi(s)$. In the literature, Ottiviani's inequality (1) is proved under the standard hypotheses (4.A) or (4.B) with the constant $\beta(s,t)$ replaced by $\beta_*(s)$.

(b): Suppose that Q is subadditive. If $(Q_1, \ldots, Q_i)$ and $Q(S_i - S_n)^*$ are independent for all $1 \le i < n$, then by (2) and (5) we see that Ottiviani's inequality (1) holds for all $s \ge 0$ and all $t \in \mathbf{R}$ with the constant $\beta(s,t)$ replaced by $\beta_*(s)$. If $X_1, \ldots, X_n$ are totally independent, then by (3), (6), (7), Thm.2.9.4 and Thm.2.10.4 we see that Ottiviani's inequality (1) holds for all $s \ge 0$ and all $t \in \mathbf{R}$ with the constant $\beta(s,t)$ replaced by $\beta_*(s)$.

(c): Suppose that $Q = \sup_{\psi\in\Psi} \psi$ for some family $\Psi \subseteq \bar{\mathbf{R}}^\Lambda$ of subadditive functions. If $X_1, \ldots, X_n$ are perfectly independent, then by (4), (5), (7), Thm.2.10.4 and Thm.2.6.2 we see that Ottiviani's inequality (1) holds for all $s \ge 0$ and all $t \in \mathbf{R}$ with the constant $\beta(s,t)$ replaced by $\beta_\Psi(s)$. In particular, we see that Ottiviani's inequality holds far beyond the standard hypotheses (4.A) or (4.B) and with a larger constant than the usual Ottaviani constant $\beta_*(s)$. If $X_1, \ldots, X_n$ are stochastic processes with parameter set T and $Q = ||\cdot||_T$ is the sup-norm, then by Def.4.7.1 we see that the functions β_* , β^* and β_Ψ take the following forms:

(10) $\beta_*(s) = \min_{1\le i<n} P_*(\|S_i - S_n\|_T \le s) \;\; , \;\; \beta^*(s) = \min_{1\le i<n} P^*(\|S_i - S_n\|_T \le s)$

(11) $\beta_\Psi(s) = \min_{1 \le i < n} \inf_{t \in T} \{ P(S_i(t) - S_n(t) \le s) \wedge P(S_n(t) - S_i(t) \le s) \}$

In particular, we see that the usual Ottaviani constant $\beta_*(s)$ is in general much smaller than $\beta_\Psi(s)$.

Proof: (1): Let us define $M^n := \max_{1 \le i \le n} Q_i$. Since $s \ge 0$, then by the definition of τ_{s+t} , we have $\{\tau_{s+t} \le n\} \subseteq \{Q_n > s+t\} \subseteq \{Q_n > t\}$ and since $\beta(s,t) \le 1$, then by the definition of $\beta(s,t)$ we have $\beta(s,t)\, P(\tau_{s+t} = i) \le P(\tau_{s+t} = i\ ,\ Q_n > t)$ for all $1 \le i \le n$. By the definition of τ_{s+t} , we have $\{\tau_{s+t} \le n\} = \{M^n > s+t\}$. So we have

$$\beta(s,t)\, P(M^n > s+t) = \beta(s,t) \sum_{i=1}^{n} P(\tau_{s+t} = i) \le \sum_{i=1}^{n} P(\tau_{s+t} = i\ ,\ Q_n > t)$$
$$\le P(Q_n > t)$$

and since Q_n is an upper P-envelope of $Q(S_n)$ and M^n is an upper P-envelope of $\max_{1 \le i \le n} Q(S_i)$ (see Lem.A.4.3), we see that (1) holds.

(2)–(4): Let $1 \le i < n$ be given. By the definition of τ_{s+t} , we have $\{\tau_{s+t} = i\} \subseteq \{Q_i > s+t\}$ and since $S_n = S_i - (S_i - S_n)$, we see that (2)–(4) follow from Lem.3.5 with $(S_1, S_2, F) := (S_i, S_i - S_n, \{\tau_{s+t} = i\}$. □

Corollary 4.7: (Lévy's inequality) *Let us consider the scenario of Thm.4.6 and let us adopt the notation of Thm.4.6. Let us define* $S_{ni} := 2S_i - S_n$ *and let* Q_{ni} *is an upper* P*-envelope of* $Q(S_{ni})$ *for all* $1 \le i \le n$ *. Let* $t \in \mathbf{R}$ *be a given number. Then Lévy's inequality:*

(1) $P^*\Big(\max_{1 \le i \le n} Q(S_i) > t \Big) \le 2\, P^*(S_n > t)$

holds in each of the following four cases

(A) Q *is a quasiconvex function satisfying*

(i) $P^*(Q_n \le t\ ,\ Q(S_{ni}) > t\ ,\ \tau_t = i) \le P^*(Q_n > t\ ,\ \tau_t = i)\ \ \forall\, 1 \le i < n$

(B) Q *is a subadditive function such that* $\{\tau_t = i\}$ *and* $\{Q(S_i - S_n)^* \le 0\}$ *are non-negatively correlated and* $P_*(Q(S_i - S_n) \le 0) \ge \frac{1}{2}$ *for all* $1 \le i < n$

(C) Q *is a subadditive function such that* $\{\tau_t = i\,, Q(S_i) > t\}$ *and* $\{Q(S_i - S_n)^* \le 0\}$ *are non-negatively correlated and* $P^*(Q(S_i - S_n) \le 0) \ge \frac{1}{2}$ *for all* $1 \le i < n$

(D) $Q = \sup_{\psi \in \Psi} \psi$ *for some non-empty set* $\Psi \subseteq \bar{\mathbf{R}}^\Lambda$ *of subadditive functions and*

(i) $(1_{\{\tau_t = i\}}, S_i)$ *and* $S_i - S_n$ *are upper* 1_{D_t}*-independent for all* $1 \le i < n$

(ii) $P^*(\psi(S_i - S_n) \le 0) \ge \frac{1}{2}\quad \forall\, 1 \le i < n\ \forall\, \psi \in \Psi$

Remarks: (a): Suppose that Q is quasiconvex and let us define $X := (X_1, \ldots, X_n)$ and $X^i := (X_1, \ldots, X_i, -X_{i+1}, \ldots, -X_n)$ for all $1 \le i < n$. By case (A), it follows easily that Lévy's inequality (1) holds for all $t \in \mathbf{R}$ in each of the following three cases. (i): The random vectors $(Q_1, \ldots, Q_i, Q_n)$ and $(Q_1, \ldots, Q_i, Q_{ni})$ have the same distribution for all $1 \le i < n$; (ii): X and X^i have the same outer distribution for all $1 \le i < n$; (iii): $(\Lambda, \mathcal{L})$ is a measurable linear space such that Q is $\mathcal{L}$-measurable, $X_1, \ldots, X_n : \Omega \to (\Lambda, \mathcal{L})$ are independent random functions and X is $\mathcal{L}$-symmetric.

(b): Suppose that Q is subadditive. By the cases (B) and (C), it follows easily that Lévy's inequality (1) holds for all $t \in \mathbf{R}$ in each of the following three cases. (i): $(Q_1, \ldots, Q_i)$ and $Q(S_i - S_n)^*$ are independent and $P_*(Q(S_i - S_n) \le 0) \ge \frac{1}{2}$ for all $1 \le i < n$; (ii): $X_1, \ldots, X_n$ are totally independent and $P^*(Q(S_i - S_n) \le 0) \ge \frac{1}{2}$ for all $1 \le i < n$; (iii) $(\Lambda, \mathcal{L})$ is a measurable linear space such that Q is $\mathcal{L}$-measurable, $X_1, \ldots, X_n : \Omega \to (\Lambda, \mathcal{L})$ are independent random functions and $P(Q(S_i - S_n) \le 0) \ge \frac{1}{2}$ for all $1 \le i < n$.

(c): Suppose that $Q = \sup_{\psi \in \Psi} \psi$ for some non-empty family $\Psi \subseteq \bar{\mathbf{R}}^\Lambda$ of subadditive functions. By case (D), it follows easily that Lévy's inequality (1) holds for all $t \in \mathbf{R}$ in each of the following two cases. (i): $X_1, \ldots, X_n$ are perfectly independent and $P^*(\psi(S_i - S_n) \le 0) \ge \frac{1}{2}$ for all $1 \le i < n$ and all $\psi \in \Psi$; (ii) $(\Lambda, \mathcal{L})$ is a measurable linear space such that Q is $\mathcal{L}$-measurable, $X_1, \ldots, X_n : \Omega \to (\Lambda, \mathcal{L})$ are independent random functions and $P(\psi(S_i - S_n) \le 0) \ge \frac{1}{2}$ for all $1 \le i < n$ and all $\psi \in \Psi$.

(d): Let $X_1, \ldots, X_n$ be stochastic processes with parameter set T and let $Q = || \cdot ||_T$ be the sup-norm. If 0 is a median of the random variable $S_i(t) - S_n(t)$ for all $t \in T$ and all $1 \le i < n$, then under the independence condition (D.i) (for instance, if $X_1, \ldots, X_n$ are perfectly independent) we have

$$(2) \qquad P^*\Big(\max_{1 \le i \le n} \|S_i\|_T > t \Big) \le 2\, P^*(\, \|S_n\|_T > t) \quad \forall\, t \in \mathbf{R}$$

Similarly, if $Q = \sup_T$ is the sup-functional and $P(S_i(t) - S_n(t) \le 0) \ge \frac{1}{2}$ for all $t \in T$ and all $1 \le i < n$, then under the independence condition (D.i) (for instance, if $X_1, \ldots, X_n)$ are perfectly independent) we have

$$(3) \qquad P^*\Big(\max_{1 \le i \le n} \sup\nolimits_T\, S_i > t \Big) \le 2\, P^*(\, \sup\nolimits_T\, S_n > t) \quad \forall\, t \in \mathbf{R}$$

(e): In the literature, Lévy's inequality (1) is proved the standard hypotheses (4.A) or (4,B) plus some strong symmetry assumptions on $X_1, \ldots, X_n$, for instance total symmetry or "symmetry" in the sense of [4; p.289]. In view of the remarks to Thm.4.6, we see that Lévy's inequality holds far beyond the standard hypotheses. Note that Lévy's inequality (1) is a consequence of Ottaviani's inequality Thm.4.6.1 (with $\beta(0,t) \ge \frac{1}{2}$). In the literature, two inequalities are proved separately under the standard hypotheses (4.A) or (4.B) with Q being a seminorm – the reason being that if Q is a seminorm, then the usual Ottavani constant $\beta_*(0) := \min_{1\le i<n} P_*(Q(S_i - S_n) = 0)$ is in general too small to be of interest (see the remarks to Thm.4.6).

(f): In the literature, there some confusion about the precise conditions for the validity of Lévy's inequality. The proofs I have found, all relies on the (implicit or explicit) assumption that $(Q_1, \dots, Q_i, Q_n)$ and $(Q_1, \dots, Q_i, Q_{ni})$ have the same distribution for all $1 \le i < n$. Consider the example in Def.3.1. Then X_1 and X_2 are independently defined stochastic processes with the same outer distribution and $X_1(t)$ and $X_2(t)$ are symmetric random variables for all $t \in T$. Hence, we see that (2) holds with $n = 2$. However, the function ξ and f may be chosen such that $||X_1 + X_2||_\Upsilon$ and $||X_1 - X_2||_\Upsilon$ are measurable but do *not* have the same distribution which means the standard proof of Lévy's maximal inequality breaks down. However, Lévy's inequality (2) do hold but for different reasons.

Proof: By Thm.4.6.1, it suffices to show that $\beta(0,t) \ge \frac{1}{2}$ in of the four cases listed in the corollary. Note that the cases (B)–(D) follow directly from Thm.4.6..2–4. So suppose that Q is a quasiconvex function satisfying (A.i) and let $1 \le i < n$ be given. Since $S_i = \frac{1}{2}(S_n + S_{ni})$ and $Q_n \ge Q(S_n)$ a.s., then by quasiconvexity of Q , we have $Q(S_i) \le Q(S_n) \vee Q(S_{ni}) \le Q_n \vee Q(S_{ni})$ a.s. and since Q_n is an upper P-envelope of $Q(S_n)$ and $\{\tau_t = i\} \subseteq \{Q_i > t\}$, then by (A.i) and Lem.A.5.5+10 we have

$$\begin{aligned} P(\tau_t = i) &= P^*(\tau_t = i \, , \, Q(S_i) > t) \\ &\le P(\tau_t = i \, , \, Q_n > t) + P^*(\tau_t = i \, , \, Q_n \le t \, , \, Q(S_{ni}) > t) \\ &\le 2\, P(\tau_t = i \, , \, Q_n > t) \end{aligned}$$

Hence, we see that $P(Q_n > t \mid \tau_t = i) \ge \frac{1}{2}$ for all $1 \le i < n$; that is, $\beta(0,t) \ge \frac{1}{2}$ and so we see that case (A) follows from Thm.4.6.1. $\square$

Theorem 4.8: (The exponential inequality) *Let Λ be a vector given space, let $Q : \Lambda \to [0,\infty]$ be a non-negative subadditive function and let $X_1, \dots, X_n : \Omega \to \Lambda$ be given functions where $n \ge 2$ is a given integer. Let $S_i := X_1 + \cdots + X_i$ denote the partial sums and let Q_i denote an upper P-envelope of $Q(S_i)$ for all $i = 1, \dots, n$. If $1 \le i < n$ is a given integer and $s, t \in \mathbf{R}$ are given numbers, we define*

$$\tau_s(\omega) := \min\{1 \le i \le n \mid Q_i(\omega) > s\} \quad \forall\, \omega \in \Omega \;\forall\, s \in \mathbf{R} \;\; (\, \min \emptyset := \infty \,)$$

$$L_n := \max_{1\le\nu\le n} Q(X_\nu) \;\; , \;\; M_n := \max_{1\le\nu\le n} Q(S_\nu) \;\; , \;\; R_i := \max_{i<\nu\le n} Q(S_\nu - S_i)$$

$$\gamma(s,t) = \max_{1\le i<n} P(R_i^* > s \mid \tau_t = i) \; , \; \gamma_\circ(s) = \max_{1\le i<n} P^*(R_i > s)$$

$$\mu(s,t) = \max_{1\le i<n} P(Q(S_n - S_i)^* > s \mid \tau_t = i) \;\; , \;\; \mu_\circ(s) = \max_{1\le i<n} P^*(Q(S_n - S_i) > s)$$

with the convention $P(F \mid G) := 0$ if $F, G \in \mathcal{F}_P$ and $P(G) = 0$ where Y^* as usual denote the (a.s. unique) upper P-envelope of $Y \in \bar{\mathbf{R}}^\Omega$. Then we have

(1) $\quad P^*(M_n > r + s + t) \le P^*(L_n > r) + \gamma(s,t)\, P^*(M_n > t) \;\; \forall\, r, s, t \ge 0$

(2) $P^*(Q(S_n) > r+s+t) \le P^*(L_n > r) + \mu(s,t)\, P^*(M_n > t) \ \ \forall\, r,s,t \ge 0$

(3) $\mu_\circ(s) \le \gamma_\circ(s) \le P^*(M_n > \frac{s}{2}) \quad \forall\, s \ge 0$

(4) *If* $(Q_1,\dots,Q_i)$ *and* R_i^* *are independent for all* $1 \le i < n$ *, then we have*

(a) $P^*(M_n > r+s+t) \le P^*(L_n > r) + \gamma_\circ(s)\, P^*(M_n > t) \ \ \forall\, r,s,t \ge 0$

(b) $P^*(M_n > r+3t) \le P^*(L_n > r) + P^*(M_n > t)^2 \ \ \forall\, r,t \ge 0$

(5) *If* $(Q_1,\dots,Q_i)$ *and* $Q(S_n - S_i)^*$ *are independent for all* $1 \le i < n$ *, then for all non-negative numbers* $r,s,t \ge 0$ *we have*

(a) $P^*(Q(S_n) > r+s+t) \le P^*(L_n > r) + \mu_\circ(s)\, P^*(M_n > t)$

(b) $P^*(Q(S_n) > r+3t) \le P^*(L_n > r) + P^*(Q(S_n) > t)\, P^*(M_n > t)$

Remarks (a): If $X_1,\dots,X_n$ are totally independent, then by Thm.2.10.3 and Thm.2.9.1 we see that the exponential inequalities (4.a) and (5.a) hold. If $(\Lambda,\mathcal{L})$ is a measurable linear space such that Q is $\mathcal{L}$-measurable and $X_1,\dots,X_n : \Omega \to (\Lambda,\mathcal{L})$ are independent random functions, we see that the exponential inequalities (4.a) and (5.a) hold. In particular, we that the exponential inequalities (4.a) and (5.a) holds far beyond the standard hypotheses (4.A) or (4,B).

(b): By (4.b), we see that under the independence hypotheses of (4) we have that the tail distribution of M_n is dominated by the tail distribution of L_n plus a term which decreases exponentially fast to zero. Let $\beta(s,t)$ denote the Ottaviani constant from Thm.4.6 and suppose that $X_1,\dots,X_n$ satisfies the independence hypothesis of (5). If $s \ge 0$ be a given number such that $\alpha := \inf_{t\ge 0} \beta(s,t) > 0$, then by (5.b) and Thm.4.6.1 we have

(6) $P^*(Q(S_n) > r+s+3t) \le P^*(L_n > r) + \frac{1}{\alpha} P^*(Q(S_n) > t)^2 \quad \forall\, r,t \ge 0$

which means that the tail distribution of $Q(S_n)$ is dominated by the tail distribution of L_n plus a term which decreases exponentially fast to zero.

Proof: Let us define $Q_0 = R_n = R_n^* := 0$, $M^0 = -\infty$ and $M^i := \max_{1\le\nu\le i} Q_\nu$ for all $1 \le i \le n$. By Lem.A.4.3, we have that M^n is an upper P-envelope of $M_n := \max_{1\le i\le n} Q(S_i)$. Since $S_\nu = X_1 + (S_\nu - S_1)$ and Q is non-negative and subadditive, we have $Q(S_\nu) \le L_n + R_1$ for all $1 \le \nu \le n$ and since $Q_0 = 0$, we have $Q_\nu \le Q_0 + L_n^* + R_1^*$ a.s. for all $1 \le \nu \le n$. Let $1 < i \le n$ be given. Since $R_n = 0$, $S_\nu = S_{i-1} + X_i + (S_\nu - S_i)$ and Q is non-negative subadditive, we have $Q(S_\nu) \le Q(S_{i-1}) + L_n + R_i$ for all $1 \le i \le \nu \le n$. Hence, we have $Q_\nu \le Q_{i-1} + L_n^* + R_i^*$ a.s. for all $1 \le i \le \nu \le n$ and since $M^0 := 0$, we have

(i) $M^n \le M^{i-1} \vee (Q_{i-1} + L_n^* + R_i^*)$ a.s. $\quad \forall\, i = 1,\dots,n$

Let us define $U_0 := 0$ and $U_i := Q(S_n - S_i)^*$ for all $1 \le i < n$. Since $Q_0 = 0$ and Q is non-negative and subadditive, then as above we obtain the following inequality:

(ii) $\quad Q_n \le Q_{i-1} + L_n^* + U_i \quad$ a.s. $\quad \forall\, i = 1, \ldots, n$

(1): Let $r, s, t \ge 0$ be given numbers and let $1 \le i \le n$ be a given integer. Since $Q_{i-1}(\omega) \le M^{i-1}(\omega) \le t$ for all $\omega \in \{\tau_t = i\}$ and $r+s+t \ge t$, then by (i) we have

$$\{\tau_t = i\,,\, M^n > r+s+t\} \subseteq \{\tau_t = i\,,\, L_n^* > r\} \cup \{\tau_t = i\,,\, R_i^* > s\}$$

Hence, we have

$$P(\tau_t = i\,,\, M^n > r+s+t) \le P(\tau_t = i\,,\, L_n^* > r) + P(\tau_t = i\,,\, R_i^* > s)$$

and since $R_n = 0$, then by Lem.A.5+10 we have

$$P(\tau_t = i\,,\, R_i^* > s) = P^*(\tau_t = i\,,\, R_i > s) \le \gamma(s,t)\, P(\tau_t = i)$$

Since $r+s+t \ge t$, we have $\{\tau_t \le n\} = \{M^n > t\} \supseteq \{M^n > r+s+t\}$ and since L_n^* and M^n are upper P-envelopes of L_n and $M_n = \max_{1 \le \nu \le n} Q(S_\nu)$, then by Lem.A.4.5 we have

$$P^*(M_n > r+s+t) = P(M^n > r+s+t) = P(\tau_t \le n\,,\, M^n > r+s+t)$$

$$= \sum_{i=1}^n P(\tau_t = i\,,\, M^n > r+s+t) \le \sum_{i=1}^n P(\tau_t = i\,,\, L_n^* > r) + \gamma(s,t) \sum_{i=1}^n P(\tau_t = i)$$

$$\le P(L_n^* > r) + \gamma(s,t)\, P(M^n > t) = P^*(L_n > r) + \gamma(s,t)\, P^*(M_n > t)$$

which proves (1).

(2): Let $r, s, t \ge 0$ be given numbers and let $1 \le i \le n$ be a given integer. Since $R^0 = 0$, then applying (ii) in the same manner as in the proof of (1), we have

$$P(\tau_t = i\,,\, Q_n > r+s+t) \le P(\tau_t = i\,,\, L_n^* > r) + P(\tau_t = i\,,\, U_i > s)$$

$$\le P(\tau_t = i\,,\, L_n^* > r) + \mu(s,t)\, P(\tau_t = i)$$

and since $\{\tau_t \le n\} = \{M^n > t\} \supseteq \{Q_n > r+s+t\}$, we see that (2) follows in the same manner as in the proof of (1).

(3): Since $R_i \ge Q(S_n - S_i)$ for all $1 \le i < n$, we see that the first inequality in (3) holds. By subadditivity of Q , we have $R_i \le 2\, M_n$ for all $1 \le i < n$ and so we see that the second inequality in (3) holds.

(4): Let $r, s, t \ge 0$ and $1 \le i < n$ be given. By the definition of τ_t , we see that $\{\tau_t = i\} \in \sigma(Q_1, \ldots, Q_i)$ and since $(Q_1, \ldots, Q_i)$ and R_i^* are independent, then Lem.A.5.5+10 and Lem.A.4.5 we have

$$P^*(\tau_t = i\,,\, R_i > s) = P(\tau_t = i\,,\, R_i^* > s) = P(\tau_t = i) \cdot P(R_i^* > s)$$

$$= P(\tau_t = i) \cdot P^*(R_i > s) \le \gamma_\circ(s)\, P(\tau_t = i)$$

Hence, we have $\gamma_\circ(s) \ge \gamma(s,t)$ and so we see that (4.a) follows from (1), and by (3) we see that (4.b) follows from (4.a) with $s := 2t$.

(5): Follows from (2) and (3) in the same manner as in the proof of (4). $\square$

Definition 4.9: (Covering and packing numbers) It is well-known that subgaussian inequalities are closely related to covering and packing numbers on semi-metric space. Recall that a *semi-metric* on the set S is a function $d : S \times S \to [0, \infty]$ satisfying $d(s,s) = 0$ and $d(s,t) = d(t,s) \le d(s,u) + d(u,t)$ for all $s,t,u \in S$.

Let (S,d) be a semi-metric space, let $A \subseteq S$ be given subset of S and let $r \ge 0$ be given non-negative number. Then we define *the d-covering number* $N_d(r,A)$ to the smallest number of closed balls of radii r and centers in A needed to cover A with the convention that $N_d(r,\emptyset) := 0$ and $N_d(r,A) := \infty$ if A cannot be covered by a finite number of closed balls of radii r with centers in A . A set $F \subseteq S$ is called *r-separated* if $d(s,t) > r$ for all $s,t \in F$ with $s \ne t$. We define *the d-packing number* to be the supremum of $|F|$ where F runs through all r-separated subsets of A and $|F|$ denotes the number of elements in F . The covering and packing numbers are related by the following well-known (and easy) inequalities:

(1) $\quad D_d(2r,A) \le N_d(r,A) \le D_d(r,A) \quad \forall\, A \subseteq S \;\forall\, r \ge 0$

Let $\operatorname{diam} A := \sup_{s,t\in A} d(s,t)$ denotes *the diameter* of A (with the convention $\operatorname{diam}\emptyset := 0$). If $A \ne \emptyset$, then we have $D_d(r,A) = N_d(r,A) = 1$ for all $r \ge \operatorname{diam} A$ and $D_d(r,A) \ge 2$ for all $0 \le r < \operatorname{diam} A$. We shall in particular consider covering and packing numbers on $\mathbf{R}^n$ with respect to the ℓ^q-norms:

(2) $\quad \|x\|_\infty := \max_{1\le i\le n} |x_i| \quad , \quad \|x\|_q := \Big\{ \sum_{i=1}^n |x_i|^q \Big\}^{\frac{1}{q}} \;\text{ if }\; 1 \le q < \infty$

If $\rho(A) := \sup_{x,y\in A} ||x-y||_\infty$ denote the $||\cdot||_\infty$-diameter of A and $\langle y \rangle$ denotes the smallest integer $\ge y$ for all $y \in \mathbf{R}$, then we have

(3) $\quad \|x\|_q \le \|x\|_p \le n^{\frac{1}{p}-\frac{1}{q}} \|x\|_q \quad \forall\, x \in \mathbf{R}^n \;\forall\, 1 \le p \le q \le \infty$

(4) $\quad N_{\|\cdot\|_q}(r,A) \le N_{\|\cdot\|_p}(r,A) \le N_{\|\cdot\|_q}\big(n^{\frac{1}{p}-\frac{1}{q}}\, r, A\big) \quad \forall\, r \ge 0 \;\forall\, 1 \le p \le q \le \infty$

(5) $\quad D_{\|\cdot\|_q}(r,A) \le D_{\|\cdot\|_p}(r,A) \le D_{\|\cdot\|_q}\big(n^{\frac{1}{p}-\frac{1}{q}}\, r, A\big) \quad \forall\, r \ge 0 \;\forall\, 1 \le p \le q \le \infty$

(6) $\quad N_{\|\cdot\|_q}(r,A) \le \Big\langle n^{\frac{1}{p}} \frac{\rho(A)}{r} \Big\rangle^n \quad \forall\, r > 0 \;\forall\, 1 \le p \le \infty$

Definition 4.10: (The subvariance) Let $X : \Omega \to \mathbf{R}$ be a random variable and let $L_X(s) := Ee^{sX}$ denote *the Laplace transform* of X . Recall that *the subvariance* of X is defined to be the smallest number $\sigma^2 \ge 0$ for which the Laplace transform L_X is dominated by the Laplace transform of the normal distribution $N(\mu,\sigma^2)$ for some $\mu \in \mathbf{R}$; that is (see [pvs.1; (4.29) p.303–304]):

(1) $\quad \gamma(X) := \inf\{\, \sigma \ge 0 \mid \exists\, \mu \in \mathbf{R} \text{ so that } L_X(s) \le e^{\mu s + \frac{1}{2}\sigma^2 s^2} \;\; \forall\, s \in \mathbf{R}\}$

If $\gamma(X) < \infty$, we say that X is *subgaussian*. If X is subgaussian, we see that the Laplace transform of X is finite on $\mathbf{R}$ and that X has finite moments of all orders. The subvariance is always greater than or equal to the ordinary variance and the subvariance shares many properties with the ordinary variance (see [7; (4.30) p.304–305] and [6; p.123–129]). Let $||X||_{\text{ex}}$ denote the Orlicz' norm of X with respect to the Orlicz' function $\text{ex}(s) := e^{s^2} - 1$. Then we have

(2) $\gamma(X) < \infty$ is and only if $||X||_{\text{ex}} < \infty$ and if so, then we have

(a) $P(X - EX > s) \leq e^{-\frac{1}{2\gamma(X)} s^2}$, $P(X - EX < -s) \leq e^{-\frac{1}{2\gamma(X)} s^2}$ $\forall s \geq 0$

(b) $P(|X - EX| > s) \leq 2\, e^{-\frac{1}{2\gamma(X)} s^2}$ $\forall s \geq 0$

(c) $P(|X| \geq s) \leq 2 \exp\Big(-\frac{s^2}{\lambda^2}\Big)$ $\forall s \geq 0 \ \forall \lambda \geq \|X\|_{\text{ex}}$

(d) $\frac{\sqrt{\gamma(X)}}{2} \leq \|X - EX\|_{\text{ex}} \leq \sqrt{6\,\gamma(X)}$

(3) If $\varepsilon = (\varepsilon_1, \dots, \varepsilon_n)$ is a Bernoulli vector, then we have

(a) $\gamma\Big(\sum_{i=1}^{n} \varepsilon_i\, x_i\Big) = \text{var}\Big(\sum_{i=1}^{n} \varepsilon_i\, x_i\Big) = \sum_{i=1}^{n} x_i^2 \ \ \forall\, (x_1, \dots, x_n) \in \mathbf{R}$

Theorem 4.11: (The Bernoulli process) *Let* $\varepsilon = (\varepsilon_1, \dots, \varepsilon_n)$ *be a Bernoulli vector and let us consider the Bernoulli process* $B(x) := \sum_{i=1}^n \varepsilon_i\, x_i$ *for* $x = (x_1, \dots, x_n) \in \mathbf{R}^n$ *. Let* $A \subseteq \mathbf{R}^n$ *be a given non-empty set, let* $N_1(r)$ *denote the* $||\cdot||_1$*-covering number of* A *, let* $D_2^\circ(r)$ *denote the* $||\cdot||_2$*-packing number of* $A^\circ := A \cup \{0\}$ *and let us define* $\theta_1(r) := 1 \vee \log N_1(r)$ *and*

$$\mathrm{B}_A := \sup_{x \in A} |B(x)| \ , \ \rho := \sup_{x,y \in A^\circ} \|x - y\|_2 \ , \ w := \sup_{x \in A} \|x\|_2 \ , \ u(r) := \theta_1(r)\, w^2$$

for all $r \geq 0$ *. Then* B_A *is* P*-measurable and we have*

(1) $P(\mathrm{B}_A > r + t) \leq 2 \exp\Big\{\log N_1(r) - \frac{t^2}{2\,w}\Big\}$ $\forall\, r, t \geq 0$

(2) $P(\mathrm{B}_A > r + t) \leq 2e \exp\Big\{-\frac{t^2}{2\,u(r)}\Big\}$ $\forall\, r, t \geq 0$

(3) $\|\mathrm{B}_A\|_{\text{ex}} \leq 16\sqrt{6} \displaystyle\int_0^\rho \sqrt{\log D_2^\circ(t)}\, dt$

(4) $P(\mathrm{B}_A > t) \leq 2 \exp\Big(-\frac{t^2}{2\,J^2}\Big)$ $\forall\, t \geq 0$ *where* $J := 8\sqrt{6} \displaystyle\int_0^\rho \sqrt{\log D_2^\circ(t)}\, dt$

Remarks: (a): Note that $B(x) = \varepsilon \bullet x$ where $(x \bullet y)$ denotes the inner product on $\mathbf{R}^n$. Hence, if $M_A(y) := \sup_{x \in A} |x \bullet y|$ denote *the Minkowski functional* of A , we have $\mathrm{B}_A = M_A(\varepsilon)$. Since M_A is a lower semicontinuous seminorm on $\mathbf{R}^n$, we see that B_A is P-measurable and if C denotes the smallest closed, symmetric, convex set containing A , then we have $M_A(y) = M_C(y)$ for all $y \in \mathbf{R}^n$ and $\mathrm{B}_A = \mathrm{B}_C = M_C(\varepsilon) = \sup_{x \in C} |\varepsilon \bullet x|$.

(b): Suppose that A is bounded, let K denote the closed convex hull of A and let A^e denote the set of extreme point of K . Then A^e is contained in the closure of A and K is the closed convex hull of A^e . Hence, we have $\mathrm{B}_A = \mathrm{B}_{A^e}$ and the inequalities (1)–(4) hold if we replace A by A^e . Since A^e is contained in the closure of A , we have that the covering and packing number of A^e is smaller than the covering and packing number of A and since A^e is contained in the relative boundary of K , we see that the covering and packing number of A^e is at least one order less that than the covering and packing number of K .

(c): The inequality (1) implies the inequality (2). The inequalities (2) and (4) are not really comparable but in some sense (4) is stronger than (2). In the literature, there exists a variety of Gaussian estimates of the tail distribution of B_A including various variants of the inequalities (1)–(4).

Proof: (1): As remarked above, we have that B_A is P-measurable. Let $r, t \geq 0$ be given and let $H(r,t)$ denote the expression on the right hand side of (1). If $w = \infty$, then A is unbounded and we have $N_1(r) = \infty$; that is, $H(s,t) = \infty$ and so we see that (1) holds. If $w = 0$, then $A = \{0\}$ and $\mathrm{B}_A = 0$ and so we see that (1) holds. If $w > 0$ and $N_1(r) = \infty$, then $H(r,t) = \infty$ and (1) holds. So suppose that $0 < w < \infty$ and $k := N_1(r, A) < \infty$. Then there exists closed $|| \cdot ||_1$-balls $B_1, \ldots, B_k$ of radii r and centers $a_1, \ldots, a_k$ belonging to A such that $A \subseteq B_1 \cup \cdots \cup B_k$. Let $x \in A$ be given. Then there exists an integer $1 \leq \nu \leq k$ such that $||x - a_\nu||_1 \leq r$. Since $B(x)$ is linear in x and $|B(x)| \leq ||x||_1$, we have

$$|B(x)| \leq |B(x - a_\nu)| + |B(a_\nu)| \leq \|x - a_\nu\|_1 + |B(a_\nu)| \leq r + |B(a_\nu)|$$

Hence, we see that $\mathrm{B}_A \leq s + \max_{1 \leq \nu \leq k} |B(a_\nu)|$ and since $a_\nu \in A$, we have $||a_\nu||_2 \leq w$ and so by Def.4.10.2+3 we have

$$P(\mathrm{B}_A \geq r + t) \leq P\Big(\max_{1 \leq \nu \leq k} |B(a_\nu)| \geq t\Big) \leq \sum_{\nu=1}^{k} P(\, |B(a_\nu)| \geq x)$$

$$\leq 2 \sum_{\nu=1}^{k} \exp\Big(-\tfrac{t^2}{2\, w^2}\Big) = H(r,t)$$

which proves (1).

(2): Let $r, t \geq 0$ be given. If $\frac{t^2}{2\, u(r)} \leq 1$, then the right hand side of (2) is ≥ 2 and so we see that (2) holds. So suppose that $\frac{t^2}{2\, u(r)} > 1$. Since $\theta_1(r) \geq 1$, we have

$$\log\, N_1(s) - \tfrac{t^2}{2\, w^2} \leq \theta_1(r) - \tfrac{t^2}{2\, w^2} = \theta_1(r) \left(1 - \tfrac{t^2}{2\, r(r)}\right) \leq 1 - \tfrac{t^2}{2\, u(r)}$$

and so we see that (2) follows from (1).

(3): By Def.4.10.2+3, we have $||B(x) - B(y)||_{\mathrm{ex}} \leq \sqrt{6}\, ||x - y||_2$ for all $x, y \in \mathbf{R}^n$. Since $0 \in A^\circ$ and $B(0) = 0$, we see that (3) follows from [15; Cor.2.2.5 p.98] (the constant $16\sqrt{6}$ is not explicitly stated in [15] but an inspection of the proof gives the constant $16\sqrt{6}$).

(4): Immediate consequence of (3) and Def.4.10.2. $\square$

Theorem 4.12: (Subgaussian inequalities) *Let T be a non-empty set and let $X_1,\dots,X_n:\Omega\to(\mathbf{R}^T,\mathcal{B}^T)$ be stochastic processes. Let $X^\epsilon=(X_1^\epsilon,\dots,X_n^\epsilon)$ denote the Bernoulli symmetrization of $X=(X_1,\dots,X_n)$ and let $S:=X_1+\cdots+X_n$ and $S^\epsilon:=X_1^\epsilon+\cdots+X_n^\epsilon$ denote the sum process and the symmetrized sum process. Let $\mu_1,\dots,\mu_n\in\mathbf{R}^T$ be given functions and let us define*

$$R_X(y,\omega):=\mathrm{B}^n(\delta\in\mathrm{E}^n\mid\|X^\epsilon(\delta,\omega)\|_T>s)\quad\forall\,y\ge 0\;\forall\,\omega\in\Omega$$

$$S_\mu(t):=\sum_{i=1}^n(X_i(t)-\mu_i(t))\;\;,\;\;\beta_\mu(s):=\inf_{t\in T}\{P(S_\mu(t)\le s)\wedge P(S_\mu(t)\ge -s)\}$$

for all $t\in T$ and all $s\ge 0$. Suppose that the following two conditions hold:

(1) $\quad ||\delta_1X_1+\cdots+\delta_nX_n||_T$ *is P-measurable for all $(\delta_1,\dots,\delta_n)\in\mathrm{E}^n$*

(2) $\quad \{\,(u_1,\dots,u_n)\in(\mathbf{R}^T)^n\mid\|u_1+\cdots+u_n\|_T>s\}\in\Sigma_P(X)\quad\forall\,s\ge 0$

Then $R_X(y,\cdot)$ is P-measurable for all $y\ge 0$ and we have

(3) $\quad \beta_\mu(s)\,P^*(\,\|S_\mu\|_T>s+2y)\le 2\,E\,R_X(y)\quad\forall\,s,y\ge 0$

Let $C,r\ge 0$ be a given non-negative numbers and let $V:\Omega\to[0,\infty]$ be a given function satisfying

(4) $\quad R_X(r+y,\omega)\le C\,\exp\Big(-\frac{y^2}{2\,V(\omega)}\Big)\quad\forall\,y>0\;\forall\,\omega\in\Omega$

If V_* is a lower P-envelope of V and Q_μ is an upper P-envelope of $||S_\mu||_T$, then for all non-negative numbers $s,y\ge 0$ we have

(5) $\quad \beta_\mu(s)\,P(\,Q_\mu\ge s+2r+2y)\le 2C\,E_*\exp\Big(-\frac{y^2}{2\,V}\Big)=2C\,E\exp\Big(-\frac{y^2}{2\,V_*}\Big)$

(6) $\quad \beta_\mu(s)\,P^*(\,Q_\mu\ge s+2r+2y)\le 2C\,P(V_*>x)+2C\,\exp\Big(-\frac{y^2}{2x}\Big)\;\;\forall\,x>0$

Let $f:[0,\infty)\to[0,\infty)$ be a increasing right continuous function such that $f(0)=0$ and let $f^\diamond(x)$ denote the following Gaussian transform of f :

(7) $\quad f^\diamond(x):=\displaystyle\int_0^\infty e^{-\frac{t^2}{2x}}\,f(dy)\quad\forall\,0\le x\le\infty$

If we define $f(\infty):=\lim_{x\to\infty}f(x)$, then we have

(8) $\quad \beta_\mu(s)\,Ef\Big(\,\tfrac{1}{2}(Q_\mu-s-2r)^+\Big)\le 2C\,Ef^\diamond(V_*)\quad\forall\,s>0$

Remarks: (a): The measurability condition (1) in Thm.4.11 goes under several names in the literature; for instance, *permissible, admissible* or *measurable*. If $\mathcal{B}$ is a Blackwell σ-algebra on T such that $X_i(t,\omega)$ is $(\mathcal{B}\otimes\mathcal{F}_P)$-measurable for all $1\le i\le n$, then the projection theorem for Blackwell spaces (see Appendix B) shows that condition (1) holds. If $X_1,\dots,X_n$ satisfies the following separability condition:

(9) There exists a countable non-empty set $D\subseteq T$ satisfying

(a) $\Big\| \sum\limits_{i=1}^{n} \delta_i\, X_i \Big\|_T = \Big\| \sum\limits_{i=1}^{n} \delta_i\, X_i \Big\|_D$ a.s. $\forall\,(\delta_1,\dots,\delta_n)\in \mathrm{E}^n$

then condition (1) trivially. Note that condition (9) holds if T is countable or more generally if T is a separable topological space such that $X_1,\dots,X_n$ have a.s. continuous sample paths. In [9; Appendix F p.262–276], you will a series of criteria implying condition (1).

(b): Recall that Thm.3.4.5–7 gives several criteria implying condition (2). Note that $P^*(||S_\mu||_T > s) = P(Q_\mu > s)\le P(Q_\mu\ge s)$. Hence, we see that (5) gives an upper estimate of the tail probability $P^*(||S_\mu||_T > s+2r+2y)$ in terms of the lower expectation of $\exp\,(-\frac{y^2}{2V})$. It is remarkable and unusual that we may estimate an upper probability in terms of a lower expectation and it means that measurability of V is not an issue here.

(c): Note that $||X^\epsilon(\,\cdot\,,\omega)||_T = \mathrm{B}_{A_\omega}$ where $\mathrm{B_A}$ is defined as in Thm.4.11 and $A_\omega := \{X(t,\omega)\mid t\in T\}\subseteq \mathbf{R}^n$ for all $\omega\in\Omega$. Hence, by (3) we see that any estimate of the upper tail distribution of the absolute maximum of the Bernoulli process gives an associated estimate of the upper tail distribution of $||S_\mu||_T$. Let $\omega\in\Omega$ be given and let d_1^ω and d_2^ω denote the following (random) semi-metric on T :

$$d_1^\omega(u,v) := \sum_{i=1}^{n} |X_i(u,\omega)-X_i(v,\omega)| = \|X(u,\omega)-X(v,\omega)\|_1 \quad \forall\, u,v\in T$$

$$d_2^\omega(u,v) := \sqrt{\sum_{i=1}^{n} |X_i(u,\omega)-X_i(v,\omega)|^2} = \|X(u,\omega)-X(v,\omega)\|_1 \quad \forall\, u,v\in T$$

Let $N_1(r,\omega)$ denote the d_1^ω-covering number of T and let $D_2(r,\omega)$ denote the d_2^ω-packing number of T . If we define

$$\theta_1(r,\omega) = 1\vee \log\, N_1(r,\omega)\ ,\ W(\omega) = \sup_{t\in T}\sqrt{\sum_{i=1}^{n} X_i(t,\omega)^2}\ ,\ V(r,\omega) = \theta_1(r,\omega)\, W(\omega)^2$$

$$\rho(\omega) = W(\omega)\vee \sup_{u,v\in T} d_2^\omega(u,v)\quad ,\quad J(\omega) = 8\,\sqrt{6}\int_0^{\rho(\omega)} \sqrt{\log\,(1+D_2(s,\omega))}\, ds$$

for all $r\ge 0$ and all $\omega\in\Omega$, then by Thm.4.11 we see that condition (4) holds with and $(C,r,V) := (2e,r,V(r))$ for all $r\ge 0$ and with $(C,r,V) := (2,0,J)$. In particular, we see that the inequalities (5), (6) and (8) holds for every fixed $r\ge 0$ with $V := V(r)$ and $C := 2e$ and that inequalities (5), (6) and (8) holds with $r := 0$ with

$V := J$ and $C := 2$. In [13; p.35–42], [15] and [4]. the inequalities (5) an (6) for these two cases has be proved more or less explicitly under the hypotheses (1), (4.A) and under some measurability assumptions on $V(r)$ or J . However, as remarked above measurability of $V(r)$ or J is *not* an issue here.

(c): Let $f : [0,\infty) \to [0,\infty)$ be a increasing right continuous function such that $f(0) = 0$ and let $f^\diamond(x)$ denote *the Gaussian transform* of f given by (7). Since $f(0) = 0$ and $\exp(-\frac{y^2}{0}) = 0$ for all $y > 0$, we see that $f^\diamond(0) = 0$ and since $\exp(-\frac{y^2}{\infty}) = 1$ for all $y \geq 0$, we see that $f^\diamond(\infty) = f(\infty)$. Note that we have the following Gaussian transforms:

$$f(x) = 1_{(\beta,\infty)}(x) \quad \Rightarrow \quad f^\diamond(x) = \exp\Big(-\frac{\beta^2}{2x}\Big)$$

$$f(x) = x^\beta \quad \Rightarrow \quad f^\diamond(x) = (2x)^{\beta/2}\,\Gamma\big(\tfrac{\beta}{2}+1\big)$$

$$f(x) = e^{\beta x} \quad \Rightarrow \quad f^\diamond(x) = \beta\,\sqrt{2\pi x}\;\, e^{\frac{1}{2}\beta^2 x}\,\Big(\tfrac{1}{2} + \operatorname{erf}(\beta\sqrt{x}\;)\Big)$$

$$f(x) = e^{\beta x^2} \quad \Rightarrow \quad f^\diamond(x) = \begin{cases} \frac{2\beta x}{1-2\beta x} & \text{if } 0 \leq x < \frac{1}{2\beta} \\[2ex] \infty & \text{if } x \geq \frac{1}{2\beta} \end{cases}$$

where $\Gamma(x)$ is the gamma function and $\operatorname{erf}(x) := (2\pi)^{-1/2}\int_0^x e^{-y^2/2}\,dy$ is error function. Hence, we see that (8) delivers various moment and exponential moment estimations of the upper P-envelope of $||S_\mu||_T$.

Proof: (3): Let us define $(\tilde\Omega,\tilde{\mathcal{F}},\tilde P) := (\mathrm{E}^n\times\Omega,\mathcal{E}^n\otimes\mathcal{F},\mathrm{B}^n\otimes P)$. Let $s,y \geq 0$ be given and let $\omega\in\Omega$ be fixed for a moment. By (1), we see that $||S^\epsilon(\delta,\omega)||_T$ is $(\mathcal{E}^n\otimes\mathcal{F}_P)$-measurable. So by Fubini's theorem, we have that $R_X(y,\cdot)$ is P-measurable and we have

$$\tilde P(\,\|S^\epsilon\|_T > s) = \int_\Omega \mathrm{B}^n(\,\|S^\epsilon(\,\cdot\,,\omega)\| > y)\,P(d\omega) = ER_X(y)$$

Let X^s denote the standard symmetrization of X and let $S^s := X_1^s + \cdots + X_n^s$ denote the symmetrized sum. Note that X^s is the standard symmetrization of $Y = (X_1-\mu_1,\ldots,X_n-\mu_n)$ and by the definition of $\Sigma_P(X)$, it follows easily that $\Sigma_P(X) = \Sigma_P(Y)$. So by (2) and Thm.4.4.4c, we have

$$\beta_\mu(s)\,P^*(\|S_\mu\|_T > s+2y) \leq 2\,\tilde P^*(\|S^\epsilon\|_T > y) \leq 2\,ER_X(y)$$

which proves (3).

(5): Let $s \geq 0$ be given. Since $v \curvearrowright \exp(-\frac{y^2}{2v})$ is increasing and continuous on $[0,\infty]$, we see that the last equality in (5) follows from Lem.A.4.4 and Lem.A.5.1. By (3), (4) and Lem.A.5.5 we have

$$\beta_\mu(s)\,P(Q_\mu > s+2r+2y) = \beta_\mu(s)\,P^*(\|S_\mu\|_T > s+2r+2y) \leq 2C\,E\exp\Big(-\frac{y^2}{2\,V_*}\Big)$$

for all $y \geq 0$ and since $y \curvearrowright \exp(-\frac{y^2}{2v})$ is decreasing, bounded and continuous on $[0,\infty)$, we see that the last expression is continuous in y. Hence, we see that (5) holds.

(6): Since $x \curvearrowright \exp(-\frac{y^2}{2x})$ is increasing on $[0,\infty]$, we see that (6) follows from (5).

(8): Let $s \geq 0$ be given and let us define $W := \frac{1}{2}(Q_\mu - s - 2r)^+$. By (5). (7), Thm.A.3.5 and Fubini's theorem, we have

$$2C\, Ef^\diamond(V_*) = 2C \int_0^\infty E\Big\{\exp\big(-\tfrac{y^2}{2\,V_*}\big)\Big\}\, f(dy)$$

$$\geq \beta_\mu(s) \int_0^\infty P^*(Q \geq s + 2r + 2y)\, f(dy) = \beta_\mu(s)\, E^* f\Big(\tfrac{1}{2}(Q - s - 2r)^+\Big)$$

which proves (8). $\square$

A. Appendix: Upper and lower expectations

This appendix contains a brief summary of the fundamental rules of calculus for upper and lower expectations and envelopes. Let $(\Omega, \mathcal{F}, P)$ be a given probability space, fixed throughout this appendix. We let $(\Omega, \mathcal{F}_P, \bar{P})$ *the completion* of $(\Omega, \mathcal{F}, P)$ and we let $\bar{M}(P)$ denote the set of all $\mathcal{F}_P$-measurable functions from Ω into the extended real line $\bar{\mathbf{R}}$. We let $L(P)$ denote the set of all random variables $X \in \bar{M}(P)$ for which the expectation EX exists but possibly equal to $\pm\infty$ and we let $L^1(P)$ denote the set of all random variables $X \in L(P)$ with finite expectation.

We extend the multiplication on $\mathbf{R}$ to the extended real line $\bar{\mathbf{R}}$ as follows

$$a \cdot \infty := \begin{cases} \infty & \text{if } 0 < a \leq \infty \\ 0 & \text{if } a = 0 \\ -\infty & \text{if } -\infty \leq a < 0 \end{cases} \quad , \quad a \cdot (-\infty) := \begin{cases} -\infty & \text{if } 0 < a \leq \infty \\ 0 & \text{if } a = 0 \\ \infty & \text{if } -\infty \leq a < 0 \end{cases}$$

and we define $\frac{1}{\pm\infty} := 0$, $\frac{1}{0} := \infty$ and $\frac{a}{b} := a \cdot \frac{1}{b}$. We extend the addition as follows $a + \infty := \infty$ if $-\infty < a \leq \infty$ and $a + (-\infty) := -\infty$ if $-\infty \leq a < \infty$. This defines $a + b$ whenever $a + b$ is not on the form $\infty + (-\infty)$ or $(-\infty) + \infty$. We shall consider the following three conventions for the remaining case:

$$\infty \,\dot{+}\, (-\infty) := \infty \;\; , \;\; \infty \underset{\cdot}{+} (-\infty) := -\infty \;\; , \;\; \infty + (-\infty) := 0$$

and we define $a \,\dot{+}\, b$ and $a \underset{\cdot}{+} b$ and $a + b$ according to these conventions. Moreover, we define $a \,\dot{-}\, b := a \,\dot{+}\, (-b)$ and $a \underset{\cdot}{-} b := a \underset{\cdot}{+} (-b)$ and $a - b := a + (-b)$. Observe that $\dot{+}$ and $\underset{\cdot}{+}$ are commutative and associative but *not* distributive and that $+$ is commutative and distributive but *not* associative.

Recall that the outer and inner probabilities P^* and P_* are given by

(A.1) $\quad P^*(A) := \inf\{\, P(F) \mid F \in \mathcal{F} \;,\; F \supseteq A\} \quad \forall\, A \subseteq \Omega$

(A.2) $\quad P_*(A) := \sup\{\, P(F) \mid F \in \mathcal{F} \;,\; F \subseteq A\} \quad \forall\, A \subseteq \Omega$

If μ is an arbitrary measure on $(\Omega, \mathcal{F})$, we define outer and inner measures $\mu^*(A)$ and $\mu_*(A)$ similarly, and we define

(A.3) $\quad \mu_\circ(A) := \sup\{\, \mu(F) \mid F \in \mathcal{F}\ ,\ \mu(F) < \infty\ ,\ F \subseteq A\} \quad \forall\, A \subseteq \Omega$

Note that $\mu_\circ \leq \mu$ with equality if and only if μ is *finitely founded* (i.e. if μ has no infinite atoms; see [7; (3.15+16) p.171–172]). Recall that E^* and E_* are given by

(A.4) $\quad \displaystyle\int^* X\,dP = E^*X := \inf\{\, EY \mid Y \in L^1(P)\ ,\ Y \geq X \ \text{ a.s.}\}$

(A.5) $\quad \displaystyle\int_* X\,dP = E_*X := \sup\{\, EY \mid Y \in L^1(P)\ ,\ Y \leq X \ \text{ a.s.}\}$

with the usual conventions $\inf \emptyset := \infty$ and $\sup \emptyset := -\infty$. If μ is an arbitrary measure on $(\Omega, \mathcal{F})$, we define $\int^* X\,d\mu$ and $\int_* X\,d\mu$ similarly.

If $X \in \bar{\mathbf{R}}^\Omega$, we let X^* and X_* denote *the upper* and *lower P-envelopes* of X; that is, the a.s. unique functions X^* and X_* satisfying

(A.6) $\quad X^* \in \bar{M}(P)$, $X^* \geq X$ a.s. and if $U \in \bar{M}(P)$ is any given random variable satisfying $U \geq X$ a.s., then we have $U \geq X^*$ a.s.

(A.7) $\quad X_* \in \bar{M}(P)$, $X_* \leq X$ a.s. and if $U \in \bar{M}(P)$ is any given random variable satisfying $U \leq X$ a.s., then we have $U \leq X_*$ a.s.

If $A \subseteq \Omega$, we let A^* and A_* denote *the P-hull* and *P-kernel of* A; that is, the a.s. unique sets A^* and A_* satisfying

(A.8) $\quad A^* \in \mathcal{F}_P \quad , \quad A^* \supseteq A$ a.s. and $\bar{P}(A^*) = P^*(A)$

(A.9) $\quad A_* \in \mathcal{F}_P \quad , \quad A_* \subseteq A$ a.s. and $\bar{P}(A_*) = P_*(A)$

Lemma A.1: (The calculus of outer and inner probabilities) *We have the following rules of calculus for outer and inner probabilities:*

(1) $\quad P^*$ *is σ-subadditive on* 2^Ω *and* P_* *is σ-superaddtive on* 2^Ω *and we have*

  (a) $\quad P_*(A^c) = 1 - P^*(A) \quad , \quad P^*(A^c) = 1 - P_*(A) \quad \forall\, A \subseteq \Omega$

(2) $\quad P_*(A) \leq P_*(A \cap B) + P^*(A \setminus B) \leq P^*(A) \quad \forall\, A, B \subseteq \Omega$

(3) $\quad \bar{P}(F) = P_*(F \cap A) + P^*(F \setminus A) \quad \forall\, F \in \mathcal{F}_P\ \forall\, A \subseteq \Omega$

(4) $\quad P_*(A) + P_*(B) \leq P_*(A \cap B) + P_*(A \cup B) \quad \forall\, A, B \subseteq \Omega$

(5) $\quad P^*(A \cap B) + P^*(A \cup B) \leq P^*(A) + P^*(B) \quad \forall\, A, B \subseteq \Omega$

(6) $\quad$ *If* $A_n \uparrow A$, *then* $P^*(A_n) \uparrow P^*(A)$

(7) *If* $A_n \downarrow A$ *, then* $P_*(A_n) \downarrow P_*(A)$

(8) *If* $A \subseteq \Omega$ *is any given set, then the set functions:*

(a) $P_A(F) := P_*(F \cap A)$ *and* $P^A(F) := P^*(F \cap A)$ *for* $F \in \mathcal{F}_P$

are finite measures on $(\Omega, \mathcal{F}_P)$ *such that* $P_A(\Omega) = P_*(A)$ *and* $P^A(\Omega) = P^*(A)$

Remark: The lemma is well-known and the proof is straight forward. □

Lemma A.2: (The calculus of upper and lower expectations) *We have the following rules of calculus for upper and lower expectations:*

(1) $E^*X = E^*X^+ \dot{-} E_*X^-$, $E_*X = E_*X^+ \underset{\cdot}{-} E^*X^-$

(2) $$E_*X \underset{\cdot}{+} E_*Y \le E_*(X \underset{\cdot}{+} Y) \le E_*(X \dot{+} Y) \le E_*X \dot{+} E^*Y \le E^*(X \dot{+} Y) \le E^*X \dot{+} E^*Y$$

(3) $E^*X = \lim_{n\to\infty} E^*(X \vee (-n))$, $E_*X = \lim_{n\to\infty} E_*(X \wedge n)$

(4) $E_*(-X) = -E^*X$, $E^*(-X) = -E_*X$

(5) $E_*(aX) = a\,E_*X$, $E^*(aX) = a\,E^*X \quad \forall\, a \in \mathbf{R}_+$

(6) $E_*(a\,1_A) = a\,P_*(A)$, $E^*(a\,1_A) = a\,P^*(A) \quad \forall\, A \subseteq \Omega\ \forall\, 0 \le a \le \infty$

(7) *If* $X_n \uparrow X$ *a.s. and* $E^*X_1 > -\infty$ *, then* $E^*X_n \uparrow E^*X$

(8) *If* $X_n \downarrow X$ *a.s. and* $E_*X_1 < \infty$ *, then* $E_*X_n \downarrow E_*X$

(9) *If* $X \ge 0$ *a.s. and* $X_n \ge 0$ *a.s. for all* $n \ge 1$ *, then we have*

(a) $E^*(\infty \cdot X) = \infty \cdot E^*X$, $E_*(\infty \cdot X) \ge \infty \cdot E_*X$

(b) $E^*\left\{\sum_{n=1}^{\infty} X_n\right\} \le \sum_{n=1}^{\infty} E^*X_n$, $E_*\left\{\sum_{n=1}^{\infty} X_n\right\} \ge \sum_{n=1}^{\infty} E_*X_n$

(10) $X \in L^1(P)$ *if and only if* $E_*X = E^*X \ne \pm\infty$ *and if* $X \in L^1(P)$ *, then we have*

(a) $E^*(aX + bY) = a\,EX + b\,E^*Y \quad \forall\, a \in \mathbf{R}\ \forall\, b \in \mathbf{R}_+$

(b) $E_*(aX + bY) = a\,EX + b\,E_*Y \quad \forall\, a \in \mathbf{R}\ \forall\, b \in \mathbf{R}_+$

(11) *If* X *is* P*-measurable, then we have*

(a) $E^*X = EX^+ \dot{-} EX^-$, $E_*X = EX^+ \underset{\cdot}{-} EX^-$

(b) $X \in L(P) \quad \Leftrightarrow \quad E^*X = E_*X$

Remark: The lemma is well-known and the proof is straight forward. Let μ be an arbitrary measure on $(\Omega, \mathcal{F})$. Then (1)–(12) hold with (E_*, E^*) replaced by $(\int_* d\mu, \int^* d\mu)$ except for the first equality in (6) which takes the following form (see (5.3))

(6*) $$\int_* a\, 1_A \, d\mu = a\, \mu_\circ(A) \quad \forall\, A \subseteq \Omega\, \forall\, 0 \le a \le \infty$$

Recall that $\mu_\circ \le \mu_*$ with equality if and only if μ is finitely founded. □

Theorem A.3: *Let μ be a measure on $(\Omega, \mathcal{F})$, let $(f_i \mid i \in I)$ be a rectilinear set of non-negative functions $f_i : \Omega \to [0, \infty]$ (see Remark (a) below) and let us define $f(\omega) := \sum_{i \in I} f_i(\omega)$ for all $\omega \in \Omega$. Then we have*

(1) *If either I is countable or if $f < \infty$ μ-a.s., then we have*

(a) $$\int^* f \, d\mu = \sum_{i \in I} \int^* f_i \, d\mu$$

(2) *If either I is finite or if $f < \infty$ μ-a.s., then we have*

(a) $$\int_* f \, d\mu = \sum_{i \in I} \int_* f_i \, d\mu$$

Let $X : \Omega \to [0, \infty]$ be a non-negative function and let $\phi : [0, \infty) \to [0, \infty)$ be an increasing right continuous function with $\phi(0) = 0$. Then we have

(3) $$\int^* X \, d\mu = \int_0^\infty \mu^*(X > x) \, dx = \int_0^\infty \mu^*(X \ge x) \, dx$$

(4) $$\int_* X \, d\mu = \int_0^\infty \mu_\circ(X > x) \, dx = \int_0^\infty \mu_\circ(X \ge x) \, dx$$

(5) $$\int^* \phi(X) \, d\mu = \int_0^\infty \mu^*(X \ge x) \, \phi(dx) \ , \quad \int_* f(X) \, d\mu = \int_0^\infty \mu_\circ(X \ge x) \, \phi(dx)$$

Remark: (a): Recall that a family $(f_i \mid i \in I)$ of extended real-valued functions $f_i : \Omega \to \bar{\mathbf{R}}$ is called *rectilinear* if the preordering $\omega \preceq \upsilon \Leftrightarrow f_i(\omega) \le f_i(\upsilon) \ \forall\, i \in I$ is *linear* (i.e. for all $\omega, \upsilon \in \Omega$ we have either $\omega \preceq \upsilon$ or $\upsilon \preceq \omega$) or equivalently if

(6) If $i \in I$ and $\omega, \upsilon \in \Omega$ are given points satisfying $f_i(\omega) < f_i(\upsilon)$, then we have $f_j(\omega) \le f_j(\upsilon)$ for all $j \in I$

Note that $(f_i \mid i \in I)$ is rectilinear if and only if $\{f_i, f_j\}$ is rectilinear for all $i, j \in I$ and if and only if the there exists a linear preordering $\le$ on Ω such that f_i is increasing on $(\Omega, \le)$ for all $i \in I$. If $(A_i \mid i \in I)$ is a family of subsets of Ω , then it follows easily that the set of indicator functions $(1_{A_i} \mid i \in I)$ is rectilinear if and only if $(A_i \mid i \in I)$ is linearly ordered by inclusion and if so, then $(a\, 1_{A_i} \mid i \in I \ , 0 \le a \le \infty)$ is rectilinear.

(b): If $(f_i \mid i \in I)$ is a rectilinear family of non-negative functions $f_i : \Omega \to [0, \infty]$ and $f(\omega) := \sum_{i \in I} f_i(\omega)$ for all $\omega \in \Omega$, then it can be shown (see [5; Thm.5.18 p.234]) that there exists a family $(\xi_i \mid i \in I)$ of increasing functions $\xi_i : \mathbf{R} \to \mathbf{R}$ satisfying

(7) $\quad \xi_i(0) = 0 \ \ \forall\, i \in I \ \ , \ \ \sum\limits_{i \in I} |\xi_i(x)| < \infty$ and $\sum\limits_{i \in I} \xi_i(x) = x \ \ \forall\, x \in \mathbf{R}$

(8) $\quad f_i(\omega) = \xi_i(f(\omega)) \quad \forall\, i \in I \ \forall\, \omega \in \{ f < \infty \}$

(9) $\quad \exists$ a countable set $J \subseteq I$ so that $\xi_i(x) = 0 \ \ \forall\, x \in \mathbf{R} \ \forall\, i \in I \setminus J$

A family $(\xi_i \mid i \in I)$ of increasing functions $\xi_i : \mathbf{R} \to \mathbf{R}$ satisfying the last two conditions in (7) is called an *increasing partition of unity*. If (ξ_i) is an increasing partition of unity, we have $\sum_{i \in I} |\xi_i(x) - \xi_i(y)| = |x - y|$ for all $x, y \in \mathbf{R}$. In particular, we see that ξ_i is increasing and Lipshitz' with Lipshitz' constant ≤ 1 for all $i \in I$.

Proof: (1): Let (ξ_i) be an increasing partition of unity satisfying (7)–(9). By (8+9), there exists a countable set $J \subseteq I$ such that $f_i(\omega) = 0$ for all $i \in I \setminus J$ and all $\omega \in \{f < \infty\}$. Hence, if $f < \infty$ μ-a.s., then $f_i = 0$ μ-a.s. for all $i \in I \setminus J$ and $f = \sum_{i \in J} f_i$ μ-a.s. Thus it suffices to prove (1) under the assumption that I is constable. Since I is at most countable, then by Lem.A.2.9 we have $\int^* f\, d\mu \leq \sum_{i \in I} \int^* f_i\, d\mu$. Hence, if $\int^* f\, d\mu = \infty$, we see that (1.a) holds. So suppose that $\int^* f\, d\mu < \infty$ and let $h \in L^1(\mu)$ be a given function satisfying $h \geq f$ μ-a.s. Let us define $\xi_i(\infty) := \infty$ and $h_i(\omega) := \xi_i(h(\omega))$ for all $i \in I$ and all $\omega \in \Omega$. By (7+8) we have $h \geq h_i \geq f_i \geq 0$ μ-a.s. for all $i \in I$. In particular, we see that $h_i \in L^1(\mu)$ and since $h < \infty$ μ-a.s., then by (7) we have $\sum_{i \in I} h_i = h$ μ-a.s. Hence, we have

$$\int_\Omega h\, d\mu = \sum_{i \in I} \int_\Omega h_i\, d\mu \geq \sum_{i \in I} \int^* f_i\, d\mu \geq \int^* f\, d\mu$$

Taking infimum over h, we obtain (1.a).

(2): If $f < \infty$ a.s., then (2.a) follows in exactly the same manner as in the proof of (1). So suppose that I is finite and let $n \geq 1$ be a given integer. Then $(f_i \wedge n \mid i \in I)$ is a rectilinear set such that $\sum_{i \in I} (f_i(\omega) \wedge n) < \infty$ for all $\omega \in \Omega$ and since $x \wedge n$ is subadditive on $[0, \infty]$, we have $f(\omega) \wedge n \leq \sum_{i \in I} (f_i(\omega) \wedge n)$ for all $\omega \in \Omega$. Hence, we have

$$\int_* (f \wedge n)\, d\mu \leq \int_* \Big\{ \sum_{i \in I} (f_i \wedge n) \Big\}\, d\mu = \sum_{i \in I} \int_* (f_i \wedge n)\, d\mu \leq \sum_{i \in I} \int_* f_i\, d\mu$$

Hence, we see that (2.a) follows from Lem.A.2.3+9.

(3): Let us define $R(x) := \mu^*(X > x)$ and $R_0(x) := \mu^*(X \geq x)$ for all $x \geq 0$. Since $R_0(x) \geq R(x) \geq R_0(y)$ for all $0 \leq x < y$, we see that $R_0(x) = R(x)$ for every continuity point of R_0 and since R_0 is decreasing, we conclude that $R_0 = R$

Lebesgue a.s. Hence, we see that the last equality in (3) holds. Let $\alpha > 1$ be a given number and let us define $X_\alpha := \sum_{n\in\mathbf{Z}} (\alpha^{n+1} - \alpha^n)\, 1_{F_n^\alpha}$ where $F_n^\alpha := \{X > \alpha^n\}$ for all $n \in \mathbf{Z} := \{0, \pm 1, \pm 2, \ldots\}$. Then it follows easily that $X \le X_\alpha \le \alpha\, X$ and since $(F_n^\alpha \mid n \in \mathbf{Z})$ is linearly ordered by inclusion, we have that $((\alpha^{n+1} - \alpha^n)\, 1_{F_n^\alpha} \mid n \in \mathbf{Z})$ is rectilinear. Since $\mu^*(F_n^\alpha) = R(\alpha^n)$, then by (1) and Lem.A.2.6 we have

$$\int^* X\, d\mu \le \int^* X_\alpha\, d\mu = \sum_{n\in\mathbf{Z}} (\alpha^{n+1} - \alpha^n)\, R(\alpha^n) \le \alpha \int^* X\, d\mu$$

and since R is decreasing, we have

$$\sum_{n\in\mathbf{Z}} (\alpha^{n+1} - \alpha^n)\, R(\alpha^n) = \sum_{n\in\mathbf{Z}} \int_{\alpha^n}^{\alpha^{n+1}} R(\alpha^n)\, dx \ge \int_0^\infty R(x)\, dx$$

$$\sum_{n\in\mathbf{Z}} (\alpha^{n+1} - \alpha^n)\, R(\alpha^n) = \alpha \sum_{n\in\mathbf{Z}} \int_{\alpha^{n-1}}^{\alpha^n} R(\alpha^n)\, dx \le \alpha \int_0^\infty R(x)\, dx$$

Hence, we have

$$\alpha^{-1} \int_0^\infty R(x)\, dx \le \int^* X\, d\mu \le \alpha^2 \int_0^\infty R(x)\, dx \quad \forall\, \alpha > 1$$

Letting $\alpha \downarrow 1$, we obtain the first inequality in (3).

(4): The last equality in (4) follows in the same manner as in the proof of (3). Let us define $R(x) := \mu_\circ(X \ge x)$ for all $x \ge 0$ and let the truncation operator $T_n : [0,\infty]^\Omega \to [0,\infty]^\Omega$ be defined as follows $T_n h := (h - \frac{1}{n})^+ \wedge (n - \frac{1}{n})$ for all $h \in [0,\infty]^\Omega$. Then it follows easily that we have $T_n h \uparrow h$ and $T_n h \le T_n g$ if $h \le g$. Let $h \in L^1(\mu)$ be a given function such that $0 \le h \le X$. Since $0 \le T_n h \le T_n X \le X$ and $T_n h \uparrow h$, then by the monotone convergence theorem, we have

$$\int_* X\, d\mu \ge \lim_{n\to\infty} \int_* T_n X\, d\mu \ge \lim_{n\to\infty} \int_\Omega T_n h\, d\mu = \int_\Omega h\, d\mu$$

Taking supremum over h , we see that $\int_* T_n X\, d\mu \to \int_* X\, d\mu$. Let $n \ge 1$ be given and let us define $F_{in} := \{X \ge \frac{i}{n}\}$ and $h_n := \frac{1}{n} \sum_{i=1}^{n^2} 1_{F_{in}})$. Then it follows easily that we have $0 \le T_n X \le h_n \le X$. Since R is decreasing with $R(\frac{i}{n}) = \mu_\circ(F_{in})$, then by (2.a) and Lem.A.2.6* we have

$$\int_* T_n X\, d\mu \le \int_* h_n\, d\mu = \tfrac{1}{n} \sum_{i=1}^{n^2} \mu_\circ(F_{in}) = \tfrac{1}{n} \sum_{i=1}^{n^2} R\big(\tfrac{i}{n}\big) \le \int_0^\infty R(x)\, dx$$

$$\int_* X\, d\mu \ge \int_* h_n\, d\mu = \tfrac{1}{n} \sum_{i=1}^{n^2} \mu_\circ(F_{in}) = \tfrac{1}{n} \sum_{i=1}^{n^2} R\big(\tfrac{i}{n}\big) \ge \int_{\frac{1}{n}}^n R(x)\, dx$$

and since $\int_* T_n X\, d\mu \to \int_* X\, d\mu$, we obtain (4) by letting $n \to \infty$.

(5): Let us define $b = \phi(\infty) := \lim_{x\uparrow\infty} \phi(x)$. If $b = 0$, then $\phi \equiv 0$ and (5) holds trivially. So suppose that $b > 0$ and let $\phi^\sim(x) := \inf\{t \ge 0 \mid \phi(t) \ge x\}$

denote the lower fractile function of ϕ for all $0 \leq x < b$. Then $\phi^\sim$ is an increasing function from $[0,b)$ into $[0,\infty)$ and since $\phi(0)=0$, we have $\phi^\sim(0)=0$. Let λ_b denote the Lebesgue measure on $[0,b)$ and let ν denote the image measure on $[0,\infty)$ of λ_b under the function $\phi^\sim$. Since ϕ is right continuous and increasing, we have $\{0 \leq x < b \mid \phi^\sim(x) \leq t\} = [0,\phi(t)]$ for all $t \geq 0$. Hence, we see that ν is the Lebesgue-Stieltjes measure induced by ϕ . Since $\{\phi(X) \geq x\} = \emptyset$ for all $x > b$ and $\{\phi(X) \geq x\} = \{X \geq \phi^\sim(x)\}$ for all $0 \leq x < b$, then by (3) we have

$$\int^* \phi(X)\, d\mu = \int_0^\infty \mu^*(\, \phi(X) \geq x)\, dx = \int_0^b \mu^*(\, \phi(X) \geq x)\, dx$$
$$= \int_0^b \mu^*(\, X \geq \phi^\sim(x))\, dx = \int_0^\infty \mu^*(X \geq y)\, \nu(dy) = \int_0^\infty \mu^*(X \geq y)\, \phi(dy)$$

which proves the first equality in (5) and the second equality follows in the same manner from (4). □

Lemma A.4: (The calculus of upper and lower envelopes I) *Let I be a non-empty finite or countable set and let $X : \Omega \to \bar{\mathbf{R}}$ and $X_i : \Omega \to \bar{\mathbf{R}}$ be given functions for all $i \in I$. Then we have the following rules of calculus of upper and lower envelopes:*

(1) $X_* = -(-X)^*$ *a.s. and* $X^* = -(-X)_*$ *a.s.*

(2) *If* $\psi : \bar{\mathbf{R}}^k \to \bar{\mathbf{R}}$ *is increasing in each variable separately, then we have*

(a) $\psi(X_1,\dots,X_k)^* \leq \psi(X_1^*,\dots,X_k^*)$ *a.s.*

(b) $\psi(X_1,\dots,X_k)_* \geq \psi(X_{1*},\dots,X_{k*})$ *a.s.*

(3) $\left(\sup\limits_{i\in I} X_i\right)^* = \sup\limits_{i\in I} X_i^*$ *a.s. and* $\left(\inf\limits_{i\in I} X_i\right)_* = \sup\limits_{i\in I} X_{i*}$ *a.s.*

(4) *If* $\phi : \bar{\mathbf{R}} \to \bar{\mathbf{R}}$ *is increasing and left continuous, then we have* $\phi(X)^* = \phi(X^*)$ *a.s. and if* $\psi : \bar{\mathbf{R}} \to \bar{\mathbf{R}}$ *is increasing and right continuous, then we have* $\psi(X)_* = \psi(X_*)$ *a.s.*

(5) $P^*(X > s) = P(X^* > s)$ *and* $P_*(X \leq s) = P(X^* \leq s)$ $\forall\, s \in \bar{\mathbf{R}}$

(6) $P_*(X \geq s) = P(X_* \geq s)$ *and* $P^*(X < s) = P(X_* < s)$ $\forall\, s \in \bar{\mathbf{R}}$

(7) *If* $U \in \bar{M}(P)$ *, then we have* $P_*(X \leq U < X^*) = 0$

(8) *If* $X \geq 0$ *a.s. and* $X^* < \infty$ *a.s. then we have* $P^*(X^* \leq \beta\, X) = 1$ *for all* $\beta > 1$

(9) *If* $U \in \bar{M}(P)$ *is a random variable satisfying* $U \geq X$ *a.s. and* $P(U > t) \leq P^*(X > t)$ *for all* $t \in \mathbf{R}$ *, then* $U = X^*$ *a.s.*

(10) *If* $U : \Omega \to [0,\infty]$ *is* P*-measurable, then we have*

(a) $U\, X_* \leq (UX)_* \leq (UX)^* \leq U\, X^*$ *a.s.*

(b) $P_*(U = \infty\,,\ X^* = 0\,,\ X < 0) = 0 \quad \Rightarrow \quad U\,X^* = (UX)^*$ *a.s.*

(c) $P_*(U = \infty\,,\ X_* = 0\,,\ X > 0) = 0 \quad \Rightarrow \quad U\,X_* = (UX)_*$ *a.s.*

Remark: The lemma is well-known and the proof is straight forward. $\square$

Lemma A.5: (The calculus of upper and lower envelopes II) *Let $U : \Omega \to [0,\infty]$ be a non-negative random variable and let $X : \Omega \to \bar{\mathbf{R}}$ be an arbitrary function. Then we have*

(1) $E_*X = E_*X_*$ *and* $E^*X = E^*X^*$

(2) $E_*(U\,X_*) \leq E_*(UX) \leq E^*(UX) \leq E^*(U\,X^*)$

(3) $P_*(U = \infty\,,\ X^* = 0\,,\ X < 0) = 0 \quad \Rightarrow \quad E^*(UX) = E^*(U\,X^*)$

(4) $P_*(U = \infty\,,\ X_* = 0\,,\ X > 0) = 0 \quad \Rightarrow \quad E_*(UX) = E_*(U\,X_*)$

(5) $\{\,X > t\}^* = \{\,X^* > t\}$ *a.s. and* $\{\,X \geq t\}_* = \{\,X_* \geq t\}\ \ \forall\, t \in \mathbf{R}$

Let I be a non-empty, finite or countable set and let A and $(A_i \mid i \in I)$ be given subsets of Ω . Then we have

(6) $A^* = \Omega \setminus (\Omega \setminus A)_*$ *a.s. and* $A_* = \Omega \setminus (\Omega \setminus A)^*$ *a.s.*

(7) $1_{A^*} = (1_A)^*$ *a.s. and* $1_{A_*} = (1_A)_*$ *a.s.*

(8) $\bigcup_{i \in I} A_i^* = \Big(\bigcup_{i \in I} A_i\Big)^*$ *a.s. and* $\bigcup_{i \in I} A_{i*} \subseteq \Big(\bigcup_{i \in I} A_i\Big)_*$ *a.s.*

(9) $\bigcap_{i \in I} A_i^* \supseteq \Big(\bigcap_{i \in I} A_i\Big)^*$ *a.s. and* $\bigcap_{i \in I} A_{i*} = \Big(\bigcap_{i \in I} A_i\Big)_*$ *a.s.*

(10) $(F \cap A)^* = F \cap A^*$ *a.s and* $(F \cup A)_* = F \cup A_*$ *a.s.* $\ \forall\, F \in \mathcal{F}_P$

Remark: The lemma is well-known and the proof is straight forward. $\square$

Theorem A.6: (The Fubini theorem I) *Let $(M_1, \mathcal{B}_1, \mu_1)$ and $(M_2, \mathcal{B}_2, \mu_2)$ be σ-finite measure spaces and let $\mu := \mu_1 \otimes \mu_2$ denote the product measure on the product space $(M_1 \times M_2, \mathcal{B}_1 \otimes \mathcal{B}_2)$. Let $h : M_1 \times M_2 \to \bar{\mathbf{R}}$ and $h_0 : M_1 \to \bar{\mathbf{R}}$ be given function. Then we have*

(1) $\displaystyle\int^* \Big\{ \int^* h(x,y)\,\mu_1(dx) \Big\}\,\mu_2(dy) \leq \int^* h(x,y)\,\mu(dx,dy)$

(2) $\displaystyle\int_* h(x,y)\,\mu(dx,dy) \leq \int_* \Big\{ \int_* h(x,y)\,\mu_1(dx) \Big\}\,\mu_2(dy)$

(3) *If* $h(x,y) = h_0(x)$ *for* μ*-a.a.* $(x,y) \in M_1 \times M_2$ *, then we have*

(a) $$\int_* h\, d\mu = \int_* h_0\, d\mu_1 \le \int^* h_0\, d\mu_1 = \int^* h\, d\mu$$

Let $f_1 : M_1 \to [0,\infty]$ *and* $f_2 : M_2 \to [0,\infty]$ *be non-negative functions and let us define* $f(x,y) := f_1(x)\, f_2(y)$ *for all* $(x,y) \in M_1 \times M_2$ *and*

$$a_i := \int_* f_i\, d\mu \ , \ p_i := (\mu_i)_*(f_i > 0) \ \textit{and} \ q_i := (\mu_i)_*(f_i = \infty) \ \textit{for} \ i = 1,2$$

Then we have

(4) $$\int^* f\, d\mu = \left\{ \int^* f_1\, d\mu_1 \right\} \cdot \left\{ \int^* f_2\, d\mu_2 \right\}$$

(5) $$\int_* f\, d\mu \ge \left\{ \int_* f_1\, d\mu_1 \right\} \cdot \left\{ \int_* f_2\, d\mu_2 \right\} \ \textit{with equality if}$$

$$0 < a_1 < \infty \ \textit{or} \ 0 < a_2 < \infty \ \textit{or} \ a_1\, a_2 > 0 \ \textit{or} \ a_1 + a_2 < \infty$$

$$\textit{or} \ p_1 = 0 \ \textit{or} \ p_2 = 0 \ \textit{or} \ a_1 < \infty \ \textit{and} \ q_2 = 0 \ \textit{or} \ a_2 < \infty \ \textit{and} \ q_1 = 0$$

Remark: The Theorem is well-known and the proof is straight forward. We may have strict inequality in (5) this is a rare exception which only occurs if all the eight cases listed in (5) fails. $\square$

Theorem A.7: (The Fubini theorem II) *Let* $(M, \mathcal{B}, \mu)$ *be a* σ*-finite measure space and let* λ *denote the Lebesgue measure on the non-negative line* $(\mathbf{R}_+, \mathcal{B}(\mathbf{R}_+))$ *. Let* $\mu \otimes \lambda$ *denote the product measure on* $M_1 \times \mathbf{R}_+$ *. Let* $\phi : M \times \mathbf{R}_+ \to [0,\infty]$ *be a given function such that* $\phi(x, \cdot)$ *is decreasing on* $[0,\infty)$ *for all* $x \in M_1$ *. Then we have*

(1) $$\int^* \phi\, d(\mu \otimes \lambda) = \int_0^\infty \left\{ \int^* \phi(x,t)\, \mu(dx) \right\} dt \ge \int^* \left\{ \int_0^\infty \phi(x,t)\, dt \right\} \mu(dx)$$

Let $(L, \mathcal{A}, \xi)$ *be a probability measure space and let* $\psi : L \times M \to [0,\infty]$ *be given function satisfying*

(2) *For every* $\varepsilon > 0$ *there exist* ξ*-measurable sets* $L_1, L_2, \ldots \in \mathcal{A}_\xi$ *and functions* $h_1, h_2, \ldots : M \to [0,\infty]$ *satisfying*

(a) $$|\psi(u,x) - \psi(v,x)| \le h_i(x) \ \ \forall\, u,v \in L_i\ \forall\, x \in M\ \forall\, i \ge 1$$

(b) $$L = \bigcup_{i=1}^\infty L_i \ \ \textit{and} \ \int^* h_i\, d\mu \le \varepsilon \ \ \forall\, i \ge 1$$

If $\eta := \xi \otimes \mu$ *denotes the product measure on* $(L \times M, \mathcal{A} \otimes \mathcal{B})$ *then we have*

(3) $$\int^* \psi\, d\eta = \int^* \left\{ \int^* \psi(u,x)\, \mu(dx) \right\} \xi(du) \ge \int^* \left\{ \int^* \psi(u,x)\, \xi(du) \right\} \mu(dx)$$

Proof: (1): Let us define $\nu := \mu \otimes \lambda$ and $\phi^\circ(t) := \int^* \phi(x,t)\,\mu(dx)$ for all $t \geq 0$. Since $\phi(x,\cdot)$ is non-negative and decreasing, we see that ϕ° is non-negative and decreasing. Let $\alpha > 1$ be given and let us define $\phi_\alpha(x,t) := \sum_{n\in\mathbf{Z}} 1_{I_n^\alpha}(t)\,\phi(x,\alpha^n)$ where $I_n^\alpha := (\alpha^n, \alpha^{n+1}]$ for all $n \in \mathbf{Z}$. Since $\phi(x,\cdot)$ is decreasing, we have $\phi \leq \phi_\alpha$ and since ϕ° is decreasing, then by Thm.A.6.4 and Lem.A.2.9 we have

$$\alpha \int_0^\infty \phi^\circ(t)\,dt \geq \alpha \sum_{n\in\mathbf{Z}} (\alpha^n - \alpha^{n-1})\,\phi^\circ(\alpha^n) = \sum_{n\in\mathbf{Z}} (\alpha^{n+1} - \alpha^n)\,\phi^\circ(\alpha^n)$$

$$= \sum_{n\in\mathbf{Z}} \int^* 1_{I_n^\alpha}(t)\,\phi(x,\alpha^n)\,\nu(dx,dt) \geq \int^* \phi_\alpha\,d\nu \geq \int^* \phi\,d\nu$$

for all $\alpha > 1$. Letting $\alpha \downarrow 1$, we see that $\int_0^\infty \phi^\circ(t)\,dt \geq \int^* \phi\,d\nu$. Thus, we see that (1) follows from Thm.A.6.1.

(3): Let $\varepsilon > 0$ be given and let (L_i, h_i) be chosen according to condition (2). By a standard disjungation, we may assume that $L_1, L_2 \ldots$ are disjoint non-empty sets and we let $a_i \in L_i$ denote an arbitrary but fixed point of L_i. Let $\psi_i(x)$ be an upper μ-envelope of $\psi(a_i,\cdot)$ for all $i \geq 1$. Let us define $h(u,x) := \sum_{i=1}^\infty 1_{L_i}(u)\,\psi(a_i,x)$ for all $i \geq 1$ and let h^* be an upper η-envelope of h. By Thm.A.8 below, we have that $1_{L_i}(u)\,\psi_i(x)$ is an upper η-envelope of $1_{L_i}(u)\,\psi(a_i,x)$ and since $\psi \geq 0$ and (L_i) is partition of L, we have $h = \sup_{i\geq 1} 1_{L_1}\,\psi(a_i,\cdot)$ and since L_i is ξ-measurable, then By Thm.A.8 below we have that $1_{L_i}(u)\,\psi_i(x)$ is an upper η-envelope of $1_{L_i}(u)\,\psi(a_i,x)$. So by Lem.A.4.10 we have

$$h^*(u,x) = \sup_{i\geq 1} \{\, 1_{L_i}(u)\,\psi_i(x)\} = \sum_{i=1}^\infty 1_{L_i}(u)\,\psi_i(x) \quad \eta\text{-a.s.}$$

Hence, by Lem.A.5.1 we have

$$\int^* h\,d\eta = \int_{L\times M} h^*\,d\eta = \sum_{i=1}^\infty \xi(L_i) \int_M \psi_i\,d\mu = \sum_{i=1}^\infty \xi(L_i) \int^* \psi(a_i,x)\,\mu(dx)$$

$$= \int^* \xi(du) \int^* h(u,x)\,\mu(dx)$$

and by (2), we have $|\psi(u,x) - h(u,x)| \leq \sum_{i=1}^\infty 1_{L_i}(u)\,h_i(x)$ for all $(u,x) \in L \times M$. Since ξ is a probability measure and (L_i) is a ξ-measurable partition of L, then by Thm.A.6.1 and Lem.A.2.9 we have

$$\int^* d\xi \int^* |\psi - h|\,d\mu \leq \int^* |\psi - h|\,d\eta \leq \sum_{i=1}^\infty \xi(L_i) \int^* h_i\,d\mu \leq \varepsilon$$

Thus, we have

$$\int^* \psi\,d\eta \leq \varepsilon + \int^* h\,d\eta = \varepsilon + \int^* d\xi \int^* h\,d\mu \leq 2\varepsilon + \int^* d\xi \int^* \psi\,d\mu$$

Letting $\varepsilon \downarrow 0$, we see that $\int^* \psi\,d\eta \leq \int^* d\xi \int^* \psi\,d\mu$ and so we see that (3) follows from Thm.A.6.1. $\square$

Theorem A.8: (Inner and outer product measures) *Let I be a non-empty countable set and let $(M_i, \mathcal{B}_i, \mu_i)$ be probability spaces for all $i \in I$. Let $\mu := \otimes_{i \in I} \mu_i$ denote the product measure on the product space $M^I := \prod_{i \in I} M_i$ equipped with its product σ-algebra $\mathcal{B}^I := \otimes_{i \in I} \mathcal{B}_i$. If $A_i \subseteq M_i$ are given sets for all $i \in I$, then we have*

(1) $$\mu_*\Bigl(\prod_{i \in I} A_i\Bigr) = \prod_{i \in I} (\mu_i)_*(A_i) \quad \textit{and} \quad \mu^*\Bigl(\prod_{i \in I} A_i\Bigr) = \prod_{i \in I} \mu_i^*(A_i)$$

Remark: The theorem is well-known, and the proofs of the first equality for a general countable sets and the second equality for finite sets I are straight forward. The proof of the second equality for an infinite countable set I is more tricky (see [1; Prop.3.6 p.I.27]; the proof is due to M. Talagrand but to the best of my knowledge, he never published it). $\square$

Definition A.9: (Iterated upper and lower expectations) Let $(M, \mathcal{B})$ be a measurable space and let $X : \Omega \to (M, \mathcal{B})$ be a given random function. Then P_X is a probability measure on $(M, \mathcal{B})$ and if $\mathcal{B}_X$ denotes the σ-algebra of all P_X-measurable subsets of M, then it follows easily that we have

(1) $X : (\Omega, \mathcal{F}_P) \to (M, \mathcal{B}_X)$ is measurable with $\bar{P}_X(B) = \bar{P}(X \in B) \;\; \forall\, B \in \mathcal{B}_X$ and if X is discrete with discrete range Q, then we have $\mathcal{B}_X = 2^M$ and if $\phi : M \to \bar{\mathbf{R}}$ is any given function satisfying $\phi(q) \geq 0$ for all $q \in Q$, then $\phi \in L(P_X)$ and we have

(a) $$E\phi(X) = \int_M \phi \, dP_X = \sum_{q \in Q} P(X = q)\, \phi(q)$$

(b) $$\bar{P}_X(A) = \bar{P}(X \in A) = \sum_{q \in Q \cap A} \bar{P}(X = q) \quad \forall\, A \subseteq M$$

We let E_X and E^X denote the upper and lower expectations with respect to the probability measure P_X; that is:

(2) $$E_X\phi := \int_* \phi(u)\, P_X(du) \quad , \quad E^X\phi := \int^* \phi(u)\, P_X(du) \quad \forall\, \phi \in \bar{\mathbf{R}}^M$$

Let $(M_1, \mathcal{B}_1)$ and $(M_2, \mathcal{B}_2)$ be a measurable spaces and let $X : \Omega \to (M_1, \mathcal{B}_1)$ and $Y : \Omega \to (M_2, \mathcal{B}_2)$ be random functions. Then we define *the upper* and *lower iterated expectations* as follows:

(3) $$E_X E_Y \psi := \int_* P_X(du) \int_* \psi(u,v)\, P_Y(dv) \quad \forall\, \psi \in \bar{\mathbf{R}}^{M_1 \times M_2}$$

(4) $$E^X E^Y \psi := \int^* P_X(du) \int^* \psi(u,v)\, P_Y(dv) \quad \forall\, \psi \in \bar{\mathbf{R}}^{M_1 \times M_2}$$

and we define $E_X\, E^Y \psi$ and $E^X\, E_Y \psi$ similarly.

Lemma A.10: *Let $(M_1, \mathcal{B}_1)$ and $(M_2, \mathcal{B}_2)$ be given measurable spaces and let $X : \Omega \to M_1$ and $Y : \Omega \to M_2$ be random functions. Let $\phi : M_1 \to \bar{\mathbf{R}}$ and $\psi : M_1 \times M_2 \to \bar{\mathbf{R}}$ be given functions. If ϕ^* and ϕ_* are upper and lower P_X-envelopes of ϕ and $\phi(X)^*$ and $\phi(X)_*$ are upper and lower P-envelopes of $\phi(X)$, then we have*

(1) $\phi_*(X) \leq \phi(X)_* \leq \phi(X)^* \leq \phi^*(X)$ *P-a.s.*

(2) $E_X\phi = E_X\phi_* = E_*\phi_*(X) \leq E_*\phi(X) = E_*\phi(X)_*$

(3) $E^X\phi = E^X\phi^* = E^*\phi^*(X) \geq E^*\phi(X) = E^*\phi(X)^*$

(4) $(P_X)_*(A) \leq P_*(X \in A) \leq P^*(X \in A) \leq P_X^*(A) \quad \forall\, A \subseteq M$

(5) *If $\psi(u,v) = \phi(u)$ for $P_{X,Y}$-a.a. $(u,v) \in M_1 \times M_2$, then we have*

  (a) $E_X\phi \leq E_{X,Y}\psi \leq E^{X,Y}\psi \leq E^X\phi$

Proof: The lemma is and easy consequence of Lem.A.2 and Lem.A.5. $\square$

Lemma A.11: *Let $(M_1, \mathcal{B}_1)$ and $(M_2, \mathcal{B}_2)$ be two given measurable spaces and let $X : \Omega \to (M_1, \mathcal{B}_1)$ and $Y : \Omega \to (M_2, \mathcal{B}_2)$ be independent random functions. Let $\phi : M_1 \to \bar{\mathbf{R}}$ $\psi : M_1 \times M_2 \to \bar{\mathbf{R}}$, $f_1 : M_1 \to [0,\infty]$ and $f_2 : M_2 \to [0,\infty]$ be given functions. If we define $f(u,v) := f_1(u)\, f_2(v)$ for all $(u,v) \in M_1 \times M_2$, then we have*

(1) $E_{X,Y}\psi \leq E_X E_Y \psi \leq E^X E^Y \psi \leq E^{X,Y}\psi$

(2) $E^{X,Y} f = E^X E^Y f = (E^X f_1) \cdot (E^Y f_2)$

(3) $(E_X f_1)\cdot(E_Y f_2) \leq E_{X,Y} f$ *with equality in each the eight cases listed in Thm.A.6.5*

(4) *If $\psi(u,v) = \phi(u)$ for $P_{X,Y}$-a.a. $(u,v) \in M_1 \times M_2$, then we have*

  (a) $E_X\phi = E_{X,Y}\psi \leq E^{X,Y}\psi = E^X\phi$

(5) *If X is discrete with discrete range $Q \subseteq M_1$, $\psi(q,v) \geq 0$ for all $q \in Q$ and all $v \in M_2$, then we have*

  (a) $E_X E^Y \psi = E^X E^Y \psi = E^{X,Y}\psi = \sum_{q \in Q} P(X = q) \cdot E^Y \psi(q, \cdot\,)$

Proof: Since $P_{X,Y} = P_X \otimes P_Y$, we see that (1)–(4) follows from Thm.A.6 and Lem.A.2.1.

(5): Let us define $\psi_u(v) := \psi(u,v)$ for all $(u,v) \in M_1 \times M_2$ and

$$\xi(u) := E^Y \psi_u \;\;,\;\; \psi_0(u,v) := \sum_{q \in Q} \psi_q(v) \cdot 1_{\{q\}}(u) \quad \forall\, u \in M_1\, \forall\, v \in M_2$$

Then we have $\psi_0(u,v)=\psi(u,v)$ for all $(u,v)\in Q\times M_2$ and $\psi_0(u,v)=0$ for all $(u,v)\notin Q\times M_2$. Since $P_X(M_1\setminus Q)=P(X\notin Q)=0$, we see that $\psi=\psi_0$ $P_{X,Y}$-a.s. and since $\psi_q(v)\geq 0$ for all $q\in Q$ and all $v\geq 0$, then by (2) and Lem.A.2.9 we have

$$E^{X,Y}\psi=E^{X,Y}\psi_0=\int^*\sum_{q\in Q}1_{\{q\}}(u)\,\psi_q(v)\,P_{X,Y}(du,dv)$$

$$\leq\sum_{q\in Q}\int^*1_{\{q\}}(u)\,\psi_q(v)\,P_{X,Y}(du,dv)=\sum_{q\in Q}(E^X1_{\{q\}})\,(E^Y\psi_q))$$

$$=\sum_{q\in Q}P(X=q)\cdot\xi(q)=E_X\xi=E_XE^Y\psi\leq E^XE^Y\psi$$

Hence, we see that (5.a) follows from (1). $\square$

B. Appendix: Perfect maps Let $(M,\mathcal{B})$ be a measurable space and let $X:\Omega\to(M,\mathcal{B})$ be a random function with distribution P_X . By Lem.A.10.4, we have $(P_X)_*(A)\leq P_*(X\in A)$ and $P^*(X\in A)\leq P_X^*(A)$ for all $A\subseteq M$. We say that X is $(P,\mathcal{B})$*-perfect* if $P^*(X\in A)=P_X^*(A)$ for all $A\subseteq M$ or equivalently if $P_*(X\in A)=(P_X)_*(A)$ for all $A\subseteq M$. In our discussion of "symmetric" random elements we need to exclude a certain possible pathological behavior of perfect functions. Let $X:\Omega\to(M,\mathcal{B})$ be random function. Then $Y(\omega_1,\omega_2):=(X(\omega_1),X(\omega_2))$ is a $(P\otimes P)$ measurable function from $\Omega\times\Omega$ into $(M\times M,\mathcal{B}\otimes\mathcal{B})$ and we say that X is *strictly* $(P,\mathcal{B})$*-perfect* if Y is $(P\otimes P,\mathcal{B}\otimes\mathcal{B})$-perfect. Since $(P\otimes P)_Y=P_X\otimes P_X$ and $Y^{-1}(A\times M)=X^{-1}(A)\times\Omega$, then by Thm.A.8 we have

(B.1) If $X:\Omega\to(M,\mathcal{B})$ is strictly $(P,\mathcal{B})$-perfect, then X is $(P,\mathcal{B})$-perfect

I don't know if the converse implication hold but I suspect that this is *not* so. Observe that we have

(B.2) If $X:\Omega\to(M,\mathcal{B})$ is (strictly) $(P,\mathcal{B})$-perfect and $Z:\Omega\to M$ is a given function such that $Z=X$ P-a.s., then $Z:\Omega\to(M,\mathcal{B})$ is (strictly) $(P,\mathcal{B})$-perfect

Let $\mathbf{N}^{\mathbf{N}}$ denote countable product of positive integers $\mathbf{N}$ and let $\mathbf{N}^{(\mathbf{N})}:=\cup_{k=1}^\infty\mathbf{N}^k$ denote the set of all *multiindices*. Recall that a *Souslin scheme* on the set M is a multiindexed family $(A(\beta)\mid\beta\in\mathbf{N}^{(\mathbf{N})})$ of subsets of M . If $A(\beta)$ is a Souslin scheme, we define *the Souslin operation* $S(A)$ and *the Souslin kernel* $\ker(A)$ as usual (see for instance [14] or [7; Exc.1.91–1.101 p.93–99]):

$$S(A):=\bigcup_{\sigma\in\mathbf{N}^{\mathbf{N}}}\bigcap_{k=1}^\infty A(\sigma_1,\ldots,\sigma_k)\quad,\quad\ker(A):=\left\{\sigma\in\mathbf{N}^{\mathbf{N}}\;\middle|\;\bigcap_{k=1}^\infty A(\sigma_1,\ldots,\sigma_k)\neq\emptyset\right\}$$

If $\mathcal{A}$ is a paving on M and $A(\beta)$ is a Souslin scheme such that $A(\beta) \in \mathcal{A}$ for all $\beta \in \mathbf{N}^{(\mathbf{N})}$, we say that $A(\beta)$ is an $\mathcal{A}$-*Souslin scheme* and we let $\mathcal{S}(\mathcal{A})$ denote the paving of all sets $B \subseteq M$ of the form $B = S(A)$ for some $\mathcal{A}$-Souslin scheme $A(\beta)$. Recall that $\mathcal{S}(\mathcal{A})$ is stable under countable unions and intersections and that $\mathcal{A} \subseteq \mathcal{S}(\mathcal{A}) = \mathcal{S}(\mathcal{S}(\mathcal{A}))$. A paving with the property that $\mathcal{A} = \mathcal{S}(\mathcal{A})$ is called *Souslin stable*. If $(M, \mathcal{B}, \mu)$ is a measure space and $\mathcal{B}_\mu$ denote the paving of all μ-measurable sets (in the sense of Caracthéodory), then it is well-known that $\mathcal{B}_\mu$ is a Souslin stable σ-algebra containing $\mathcal{B}$. If we let $\mathcal{B}_u$ denote the paving of *universally measurable* subsets of M; that is, $\mathcal{B}_u := \cap_\mu \mathcal{B}_\mu$ where μ ranges over all measures on $(M, \mathcal{B})$, then $\mathcal{B}_u$ is a Souslin stable σ-algebra containing $\mathcal{B}$.

Let T be a Hausdorff space. Then we say that T is *Polish* if T is separable and the topology is generated by a complete metric, we say that T is *analytic* if T is a continuous image of some Polish space, and we say that T is *K-analytic* if T is *usco* image of a Polish (i.e. there exists a Polish space S and a map $\theta : S \to 2^T$ such that $T- = \cup_{s \in S} \theta(s)$, the set $\theta(s)$ is compact for all $s \in S$ and $\{s \in S \mid \theta(s) \cap F \neq \emptyset\}$ is closed for every closed set $F \subseteq T$; see [14]). Let $(M, \mathcal{B})$ be a given measurable space. Recall that $(M, \mathcal{B})$ is called

- *Blackwell* if $f(M)$ is an analytic subset of $\mathbf{R}$ for every $\mathcal{B}$-measurable function $f : M \to \mathbf{R}$ or equivalently if $\ker(A)$ is an analytic subset of $\mathbf{N}^{\mathbf{N}}$ for every $\mathcal{B}$-Souslin scheme $A(\beta)$
- *discrete* if there exists a family $(D_i \mid i \in I)$ of mutually disjoint subsets of M such that $\mathcal{B} = \sigma(D_i \mid i \in I)$
- σ-*discrete* if there exists a finite or countable family $(D_i \mid i \in I)$ of mutually disjoint subsets of M such that $\mathcal{B} = \sigma(D_i \mid i \in I)$
- *countably separated* if there exists an injective, real-valued, $\mathcal{B}$-measurable function $f : M \to \mathbf{R}$
- *countably generated* if the σ-algebra $\mathcal{B}$ is generated by a countable paving

If $(S, \mathcal{A})$ and $(M, \mathcal{B})$ are measurable spaces and $f : (M.\mathcal{B}) \to (S, \mathcal{A})$ is a measurable function, then we have the following stability properties for Blackwell spaces:

- **The measurable image theorem:** Let $(M, \mathcal{B})$ be a Blackwell space and let $L \in \mathcal{S}(\mathcal{B})$ be a given $\mathcal{B}$-Souslin set. If $\mathcal{B}_L := \{L \cap B \mid B \in \mathcal{B}\}$ and $\mathcal{A}_{f(L)} := \{f(L) \cap A \mid A \in \mathcal{A}\}$ denote the trace σ-algebras on L and $f(L)$, respectively, then we have
  - (a) $(L, \mathcal{B}_L)$ and $(f(L), \mathcal{A}_{f(L)})$ are Blackwell spaces
  - (b) If $(S, \mathcal{A})$ is countably separated, then $f(L) \in \mathcal{S}(\mathcal{A}) \subseteq \mathcal{A}_u$ and f is strictly $(\mu, \mathcal{A})$-perfect for every probability measure μ on $(M, \mathcal{M})$

- **The inverse measurable image theorem:** If $(S, \mathcal{A})$ is a Blackwell space and $\mathcal{B} = f^{-1}(\mathcal{A})$, then $(M, \mathcal{B})$ is a Blackwell space.

- **The product theorem:** If $(M_i, \mathcal{B}_i \mid i \in I)$ is a an arbitrary family of Blackwell spaces, then the product space $(\prod_{i\in I} M_i, \otimes_{i\in I} \mathcal{B}_i)$ is a Blackwell space.

- **Analytic spaces:** If M is an analytic space and $\mathcal{B}$ is the Borel σ-algebra on M (i.e., the smallest σ-algebra containing all open sets), then $(M, \mathcal{B})$ is a Blackwell space.

- **K-Analytic spaces:** If M is a K-analytic space and $\mathcal{B}$ is the Baire σ-algebra on M (i.e., the smallest σ-algebra making all continuous real-valued functions measurable), then $(M, \mathcal{B})$ is a Blackwell space.

- **Discrete spaces:** If $(M, \mathcal{B})$ is a discrete measurable space, then $(M, \mathcal{B})$ is a is Blackwell space.

- **The projection theorem:** Let $\pi(s, x) := s$ denotes the projection of $S \times M$ onto S . If $(M, \mathcal{B})$ is a Blackwell space and $C \in \mathcal{S}(\mathcal{A} \otimes \mathcal{B})$, then $\pi(C) \in \mathcal{S}(\mathcal{A}) \subseteq \mathcal{A}_u$.

- **The measurable selection theorem:** Let $\pi(s, x) := s$ denotes the projection of $S \times M$ onto S . If $(M, \mathcal{B})$ is a countably generated Blackwell space and $C \in \mathcal{S}(\mathcal{A} \otimes \mathcal{B})$, then there exists a measurable function $h : (S, \mathcal{A}_u) \to (M, \mathcal{B})$ such that $(s, h(s)) \in C$ for all $s \in \pi(C)$.
  Remark: The measurable selection theorem, fails in general for arbitrary Blackwell spaces but it do hold for a much larger class than the countably generated Blackwell space; for instance, if M is a K-analytic space and $\mathcal{B}$ is the Baire σ-algebra on $\mathcal{B}$, of if $(M, \mathcal{B}) = (\mathbf{R}^T, \mathcal{B}^T)$ for some non-empty set T .

Lemma B.1: *Let $(M, \mathcal{B})$ be a measurable space and let $X, Y : \Omega \to (M, \mathcal{B})$ be two $(P, \mathcal{B})$-perfect functions. Let $\phi : M \to \bar{\mathbf{R}}$ be a given function. If ϕ^* and ϕ_* are any given upper and lower P_X-envelopes of ϕ and $\phi(X)^*$ and $\phi(X)_*$ are any given upper and lower P-envelopes of $\phi(X)$, then we have*

(1) $X(\Omega) \in \mathcal{B}_X$ *and we have* $B \in \mathcal{B}_X$ *if and only if* $X^{-1}(B) \in \mathcal{F}_P$

(2) *If X and Y have the same distribution on $(M, \mathcal{B})$, then X and Y have the same outer distribution*

(3) $\phi^*(X) = \phi(X)^*$ *P-a.s. and* $\phi_*(X) = \phi(X)_*$ *P-a.s.*

(4) $E_X\phi = E_*\phi(X)$ *and* $E^X\phi = E^*\phi(X)$

Proof: (1) and (2) follows easily from the definition of perfectness. Since X is perfect, we have $P^*(\phi(X) > t) = P_X^*(\phi > t)$ and since $\phi^*(X)$ is a P-measurable function such that $\phi^*(X) \geq \phi(X)$ P-a.s., then by Lem.A.4.9 we have $\phi^*(X) = \phi(X)^*$

P-a.s. and in the same manner, we see that $\phi_*(X) = \phi(X)_*$ P-a.s. Hence, we see that (3) holds and (4) follows from (3) and Lem.A.5.1. $\square$

Lemma B.2: *Let $(M, \mathcal{B})$ be a measurable space and let $X : \Omega \to M$ be measurable function. Let $x \curvearrowright P(F \mid X = x)$ denote a given and fixed version of the conditional probability for any given set $F \in \mathcal{F}$. Then the following eight statements are equivalent:*

(1) X *is $(P, \mathcal{B})$-perfect*

(2) X *is $(\bar{P}, \mathcal{B}_X)$-perfect*

(3) $P_*(X^{-1}(A)) = 0 \;\; \forall\, A \subseteq M$ *satisfying* $(P_X)_*(A) = 0$

(4) $\forall\, A \subseteq M \;\exists\, B \in \mathcal{B}$ *so that* $B \supseteq A$ *and* $P_*(X^{-1}(B \setminus A)) = 0$

(5) $\forall\, A \subseteq M \;\exists\, B \in \mathcal{B}_X$ *so that* $B \supseteq A$ *and* $P_*(X^{-1}(B \setminus A)) = 0$

(6) $\forall\, F \in \mathcal{F}_P \;\exists\, K \in \mathcal{F}_P$ *so that* $K \subseteq F$, $\bar{P}(F \setminus K) = 0$ *and* $X(K) \in \mathcal{B}$

(7) $\forall\, F \in \mathcal{F}$ *with* $P(F) > 0 \;\exists\, K \subseteq F$ *so that* $P^*(K) > 0$ *and* $X(K) \in \mathcal{B}_X$

(8) $P(F \mid X = x) = 0$ *for P_X-a.a.* $x \in M \setminus X(F) \;\; \forall\, F \in \mathcal{F}$

Proof: (1) $\Leftrightarrow$ (2): Since $\bar{P}^* = P^*$, we see that the equivalence of (1) and (2) follows from Def.A.9.1.

(1) $\Rightarrow$ (3): Evident!

(3) $\Rightarrow$ (4): Suppose that (3) holds and let $A \subseteq M$ be given. Let $B \in \mathcal{B}$ be chosen such that such that $B \supseteq A$ and $P_X^*(A) = P_X(B)$. Since B is a P_X-hull of A , we have $(P_X)_*(B \setminus A) = 0$ and so by (3), we have $P_*(X^{-1}(B \setminus A)) = (P_X)_*(B \setminus A) = 0$ which proves (4).

(4) $\Rightarrow$ (5): Evident!

(5) $\Rightarrow$ (6): Suppose that (5) holds and let $F \in \mathcal{F}_P$ be given. Let us define $A := X(F)$ and $A_0 := M \setminus A$. By (5), there exists $B_0 \in \mathcal{B}_X$ such that $B_0 \supseteq A_0$ and $P_*(X^{-1}(B_0 \setminus A_0)) = 0$ and since $B_0 \in \mathcal{B}_X$, there exists $B_1 \in \mathcal{B}$ such that $B_1 \supseteq B_0$ and $\bar{P}_X(B_1 \setminus B_0) = 0$. Hence, we see that $X^{-1}(B_1 \setminus B_0)$ is a P-null set and consequently, we have $P_*(X^{-1}(B_1 \setminus A)) = 0$. Let us define $B := M \setminus B_1$ and $K := F \cap X^{-1}(B)$. Then we have $B \in \mathcal{B}$, $K \in \mathcal{F}_P$, $K \subseteq F$ and $P_*(X^{-1}(A \setminus B)) = 0$. Since $F \subseteq X^{-1}(A)$, we have $F \setminus K \subseteq X^{-1}(A \setminus B)$ and since $F \setminus K \in \mathcal{F}_P$, we see that $\bar{P}(F \setminus K) = 0$. Since $B \subseteq X(F)$, we have $X(K) = X(F \cap X^{-1}(B)) = B \in \mathcal{B}$ and so we see that K have properties listed in (6).

(6) $\Rightarrow$ (7): Evident!

(7) $\Rightarrow$ (1): Suppose that (7) holds and let $A \subseteq M$ be given. Let $B \in \mathcal{B}$ be chosen such that $B \subseteq A$ and $P_X(B) = P_X^*(A)$. Suppose that $P_*(X^{-1}(B \setminus A)) > 0$. Then there exists $F \in \mathcal{F}$ such that $P(F) > 0$ and $F \subseteq X^{-1}(B \setminus A)$. So by (6) there exists a set $K \subseteq F$ such that $P^*(K) > 0$ and $X(K) \in \mathcal{B}_X$. Since

$X(K) \in \mathcal{B}_X$ and $X(K) \subseteq X(F) \subseteq B \setminus A$, we have $X^{-1}(X(K)) \in \mathcal{F}_P$ and $\bar{P}(X^{-1}(X(K))) = \bar{P}_X(X(K)) = 0$ and since $K \subseteq X^{-1}(X(K)) \in \mathcal{F}$, we have $P^*(K) = 0 < P^*(K)$ which is impossible. Thus, we have $P_*(X^{-1}(B \setminus A)) = 0$ and since $A \subseteq B$ and $X^{-1}(B) \in \mathcal{F}_P$, then by Lem.A.1.3 we have

$$P_X^*(A) = P_X(B) = \bar{P}(X^{-1}(B)) = P^*(X^{-1}(A)) + P_*(X^{-1}(B \setminus A)) = P^*(X^{-1}(A))$$

Thus, we see that X is $(P, \mathcal{B})$-perfect.

(1) $\Rightarrow$ (8): Suppose that X is $(P, \mathcal{B})$-perfect and let $F \in \mathcal{F}$ be given. Let $B \in \mathcal{B}$ be a lower P_X-envelope of $A := X(F)$ such that $B \subseteq A$. Since X is $(P, \mathcal{B})$-perfect, then by Lem.B.1.3 we have $X^{-1}(B)$ is a lower P-envelope of $X^{-1}(X(F))$ and since $F \subseteq X^{-1}(A)$, we have $F \subseteq X^{-1}(B)$ a.s. Hence, we have

$$P(F \mid X = x) \leq P(X^{-1}(B) \mid X = x) = 1_B(x) \quad P_X\text{-a.s.}$$

and since $B \subseteq A = X(F)$, we see that (8) holds.

(8) $\Rightarrow$ (1): Suppose that (8) holds and let $A \subseteq M$ be a given set. Let $F \in \mathcal{F}$ be a lower P-envelope of $X^{-1}(A)$ such that $F \subseteq X^{-1}(A)$. Then the set $C := \{x \in M \mid P(F \mid X = x) > 0\}$ belongs to $\mathcal{B}$ and we have

$$P_*(X^{-1}(A)) = P(F) = \int_C P(F \mid X = x)\, P_X(dx)$$

By (8), we have $C \subseteq X(F)$ P_X-a.s. and since $P(F \mid X = x) \leq 1$ P_X-a.s., we have $P_*(X^{-1}(A) \leq (P_X)_*(X(F)))$. Since $F \subseteq X^{-1}(A)$, we have $X(F) = X(\Omega) \cap A \subseteq A$, and so by Lem.A.10.4 we conclude that $P_*(X^{-1}(A) = (P_X)_*(A)$ for all $A \subseteq M$; that is, X is $(P, \mathcal{B})$-perfect. $\square$

Definition B.3: (The useful rule) Let $X : \Omega \to (M, \mathcal{B})$ be a random function and let $F \in \mathcal{F}_P$ be a given set. Recall that *the useful rule* states, loosely speaking, that we may use what we know in the computation of the conditional probability $P(F \mid X = u)$; that is, $P(F \cap \{X = u\} \mid X = u) = P(F \mid X = u)$. Since the conditional probability $P(F \mid X = u)$ only is defined P_X-a.s. (and since $\{X = u\}$ need not be P-measurable) this formulation of the useful rule is meaningless. Since $F \cap X^{-1}(u) = \emptyset$ if and only if $u \notin X(F)$, we see that the property (8) in Lem.B.2 offers a possible (rigorous) definition of the useful rule. There are several other formulations of the useful rule – the following is the most common:

Let $X : \Omega \to (M, \mathcal{B})$ be a random function. Recall that a *regular conditional probability of* P *given* X is map $Q(F \mid x)$ from $\mathcal{F} \times M$ into $[0, 1]$ satisfying

(1) $Q(\cdot \mid x)$ is a probability measure on $(\Omega, \mathcal{F})$ for all $x \in M$ and $Q(F \mid \cdot)$ is P-measurable for all $F \in \mathcal{F}$

(2) $\displaystyle\int_B Q(F \mid x)\, P_X(dx) = \bar{P}(F \cap X^{-1}(B)) \quad \forall\, B \in \mathcal{B}\ \forall\, F \in \mathcal{F}$

Note that (2) just states that $Q(F \mid x)$ is a version of the conditional probability $P(F \mid X = x)$. If Q is a regular conditional probability of P given X , we say that Q is a *useful regular conditional probability of* P *given* X if the following form of *the useful rule* holds:

(3) $Q^*(X^{-1}(x) \mid x) = 1$ for P_X-a.a. $x \in M$

where $Q^*(A \mid x)$ denotes the upper $Q(\,\cdot\mid x)$-probability of $A \subseteq \Omega$. If Q is a useful regular conditional probability of P given X and $N \in \mathcal{B}$ is the exceptional null set from (3), then we have the following strong form of the useful rule:

(4) $Q^*(F \cap \{ X = x\} \mid x) = 1 \ \ \forall x \in M \setminus N \ \forall F \in \mathcal{F}$

A slight modification of the proof of [pvs.2; (10.29) p.155] gives the following existence theorem for regular conditional probabilities:

Theorem B.4: *Let* $(M, \mathcal{B})$ *be a measurable space and let* $X : \Omega \to (M, \mathcal{B})$ *be a random function. If there exists a semicompact paving* $\mathcal{K} \subseteq \mathcal{F}$ *such that* $\mathcal{K}$ *is stable under finite unions and countable intersections and*

(1) $P(F) = \sup\{ P(K) \mid K \in \mathcal{K} \ , \ K \subseteq F\} \quad \forall F \in \mathcal{F}$

(2) $\emptyset \in \mathcal{K} \subseteq \mathcal{B} \subseteq \mathcal{S}$ *where* $\mathcal{S}$ *denotes the smallest Souslin closed* σ*-algebra on* M *containing the paving* $\widehat{\mathcal{K}}$ *given by*

(a) $\widehat{\mathcal{K}} := \{ C \subseteq M \mid C \cap K \in \mathcal{K} \ \ \forall K \in \mathcal{K}\}$

Then there exists a regular conditional probability Q *of* P *given* X *satisfying*

(3) $Q(F \mid x) = \sup\{ Q(K \mid x) \mid K \in \mathcal{K} \ , \ K \subseteq F\} \quad \forall F \in \mathcal{F} \, \forall x \in M$

Remark: Recall that a paving $\mathcal{K}$ is *semicompact* if for every sequence $K_1, K_2, \ldots \in \mathcal{K}$ with $\cap_{i=1}^{\infty} K_i = \emptyset$, there exists an integer $n \geq 1$ such that $\cap_{i=1}^{n} K_i = \emptyset$. Let T be a non-empty set and let $\pi_\alpha : \mathbf{R}^T \to \mathbf{R}^\alpha$ denote the projection whenever $\alpha \subseteq T$ is a non-empty set. Let $\mathcal{K}$ denote the paving of all subsets of $\mathbf{R}^T$ of the form $\pi_\alpha^{-1}(C)$ where $\alpha \subseteq T$ is a non-empty countable set and $C \subseteq \mathbf{R}^\alpha$ is a compact set in the product topology on $\mathbf{R}^\alpha$. Then we have

(4) $\mathcal{K}$ is a semicompact paving such that $\mathcal{K}$ is stable under finite unions and countable intersections and $\emptyset \in \mathcal{K} \subseteq \mathcal{B}^T = \sigma(\mathcal{K})$

(5) If μ is any given probability measure on $(\mathbf{R}^T, \mathcal{B}^T)$, then we have

(a) $\mu(B) = \sup\{ \mu(K) \mid K \in \mathcal{K} \ , \ K \subseteq B\} \quad \forall B \in \mathcal{B}^T$

Hence, if the underlying probability space $(\Omega, \mathcal{F}, P)$ is of the form $(\Omega, \mathcal{F}) = (\mathbf{R}^T, \mathcal{B}^T.P)$ for some non-empty set T and some probability measure P on $(\mathbf{R}^T, \mathcal{B}^T)$ (which by Kolmogorov's consistency theorem may be assumed in most cases), then any random function admits a regular conditional probability.

Theorem B.5: *Let $(M, \mathcal{B})$ be a measurable space and let $X : \Omega \to (M, \mathcal{B})$ be a random function. Let $Q(F \mid x)$ be regular conditional probability of P given X and let us define $Y(\omega) := (\omega, X(\omega))$ for all $\omega \in \Omega$. Then $Y : \Omega \to (\Omega \times M, \mathcal{F} \otimes \mathcal{B})$ is a random function and we have*

(1) $\mathcal{F}_P \otimes \mathcal{B}_X \subseteq (\mathcal{F} \otimes \mathcal{B})_Y$

Let $H \subseteq \Omega \times M$ be a given P_Y-measurable set and let $H_x := \{\omega \in \Omega \mid (\omega, x) \in H\}$ denote the x-section of H for all $x \in M$. Then we have

(2) *The functions $x \curvearrowright Q^*(H_x \mid x)$ and $x \curvearrowright Q_*(H_x \mid x)$ are P_X-measurable and we have $Q^*(H_x \mid x) = Q_*(H_x \mid x)$ for P_X-a.a. $x \in M$*

(3) $\bar{P}_Y(H) = \int_M Q^*(H_x \mid x)\, P_X(dx) = \int_M Q_*(H_x \mid x)\, P_X(dx)$

(4) *If Q is a useful regular conditional probability of P given X , then X is strictly $(P, \mathcal{B})$-perfect*

(5) *If $Y(\Omega)$ is P_Y-measurable, then Q is useful*

(6) *If Y is $(P, \mathcal{F} \otimes \mathcal{B})$-perfect, then $Y(\Omega)$ is P_Y-measurable and Q is useful*

(7) *If $(M, \mathcal{B})$ is countably separated, then $Y(\Omega) \in \mathcal{F}_P \otimes \mathcal{B}$ and Q is useful*

Proof: (1): Since P and P_X are the marginal distributions of $Y(\omega) = (\omega, X(\omega))$, it follows easily that (1) holds.

(2+3): Let $\mathcal{H}$ denote the paving of all sets $H \in \mathcal{H}_Y$ satisfying (2+3). Then a standard argument shows that $\mathcal{H}$ is stable under complementation, countable increasing unions and countable decreasing intersections and that $F \times B \in \mathcal{H}$ for all $F \in \mathcal{F}$ and all $B \in \mathcal{B}$. So by Sierpińsky's lemma (see [7. (1.6.3) p.9]) we conclude that $\mathcal{F} \otimes \mathcal{B} \subseteq \mathcal{H}$. Let $H \in (\mathcal{F} \otimes \mathcal{B})_Y$ be given. Then there exist sets $F, G \in \mathcal{F} \otimes \mathcal{B}$ such that $F \subseteq H \subseteq G$ and $P_Y(G \setminus F) = 0$. Since $G \setminus F \in \mathcal{F} \otimes \mathcal{B} \subseteq \mathcal{H}$, we have

$$0 = P_Y(G \setminus F) = \int_M Q(G_x \setminus F_x \mid x)\, P_X(dx)$$

Hence, there exists a P_X-null set N such that $Q(G_x \setminus F_x \mid x) = 0$ for all $x \in M \setminus N$ and since $F_x \subseteq H_x \subseteq G_x$, we see that H_x is $Q(\cdot \mid x)$-measurable with $Q^*(H_x \mid x) = Q_*(H_x \mid x) = Q(G_x \mid x)$ for all $x \in M \setminus N$. Since $G \in \mathcal{H}$ and $P_Y(G) = \bar{P}_Y(H)$, we see that $H \in \mathcal{H}$ for all $H \in (\mathcal{F} \otimes \mathcal{B})_Y$ which proves (2) and (3).

(4): Suppose that $Q(F \mid x)$ is a useful regular conditional probability of P given X . Let us define $Z(\omega_1, \omega_2) := (X(\omega_1), X(\omega_2))$ for all $(\omega_1, \omega_2) \in \Omega \times \Omega$. Then it follows easily that $R(\,\cdot \mid x, y) := Q(\,\cdot \mid x) \otimes Q(\,\cdot \mid y)$ defines a regular distribution of $P \otimes P$ given Z and since Q is useful and $P_Z = P_X \otimes P_X$, then by Thm.A.8 we have

$$R^*(Z^{-1}(x,y) \mid x,y) = R^*(X^{-1}(x) \times X^{-1}(y) \mid x,y)$$

$$= Q^*(X^{-1}(x) \mid x) \cdot Q^*(X^{-1}(y) \mid y) = 1 \ \ P_Z\text{-a.s.}$$

Hence, we see that R is a useful regular probability of P given Z . So by Def.B.3.4 and Lem.B.2.1+8 we conclude that Z is $(P \otimes P, \mathcal{B} \otimes \mathcal{B})$-perfect; that is, X is strictly $(P, \mathcal{B})$-perfect.

(5): Suppose that $Y(\Omega) \in (\mathcal{F} \otimes \mathcal{B})_y$. Since $P(Y \in H) = 1$ for all sets $H \in \mathcal{F} \otimes \mathcal{B}$ satisfying $H \supseteq Y(\Omega)$, we have $\bar{P}_Y(Y(\Omega)) = P_Y^*(Y(\Omega)) = 1$ and since $X^{-1}(x) = \{\omega \mid (\omega, x) \in Y(\Omega)\}$ for all $x \in M$, then by (3) we have $Q^*(X^{-1}(x) \mid x) = 1$ for P_X-a.a. $x \in M$; that is, Q is a useful regular probability.

(6) follows from (5) and Lem.B.1.1, and since $Y(\Omega)$ is the graph of X , we see that (7) follows from (5) and [7; Exc.1.89 p.93]. ☐

Lemma B.6: *Let $(M_1, \mathcal{B}_1)$ and $(M_2, \mathcal{B}_2)$ be given measurable spaces and let $X_1 : \Omega \to M_1$ and $X_2 : \Omega \to M_2$ be random functions. Let $\xi : M_1 \to (M_2, \mathcal{B}_2)$ be a P_X-measurable function. Then the composite function $Y = \xi \circ X : \Omega \to (M_2, \mathcal{B}_2)$ is a random function and we have*

(1) *If X is $(P, \mathcal{B}_1)$-perfect and ξ is $(P_X, \mathcal{B}_2)$-perfect, then Y is $(P, \mathcal{B}_2)$-perfect*

(2) *If X is strictly $(P, \mathcal{B}_1)$-perfect and ξ is strictly $(P_X, \mathcal{B}_2)$-perfect, then Y is $(P, \mathcal{B}_2)$-perfect*

(3) *If Y is $(P, \mathcal{B}_2)$-perfect, then ξ is $(P_X, \mathcal{B}_2)$-perfect*

(4) *If Y is strictly $(P, \mathcal{B}_2)$-perfect, then ξ is strictly $(P_X, \mathcal{B}_2)$-perfect*

(5) *If Y is $(P, \mathcal{B}_2)$-perfect and there exists a P_X-null set $N \subseteq M_1$ such that the restriction of ξ to the set $M_1 \setminus N$ is injective, then X is $(P, \mathcal{B}_1)$-perfect*

(6) *If Y is strictly $(P, \mathcal{B}_2)$-perfect and there exists a P_X-null set $N \subseteq M_1$ such that the restriction of ξ to the set $M_1 \setminus N$ is injective, then X is strictly $(P, \mathcal{B}_1)$-perfect*

(7) *If $\eta : M_1 \to (\Omega, \mathcal{F})$ is a P_X-measurable function such that $\eta(X(\omega)) = \omega$ P-a.s. and there exists aP_X-null set $N \subseteq M_1$ such that the restriction of η to $M_1 \setminus N$ is injective, then X is strictly $(P, \mathcal{B}_1)$-perfect*

(8) *If X is a bijection of Ω onto L and its inverse $X^{-1} : L \to (\Omega, \mathcal{F})$ is P_X-measurable, then X is strictly $(P, \mathcal{B}_1)$-perfect*

Proof: Let us define $Q = P_X$. By Def.A.9.1, we see that $Y : \Omega \to (M_2, \mathcal{B}_2)$ is a random function and observe that we have $P_Y = Q_\xi$.

(1): Since $P_Y = Q_\xi$, we see that (1) holds.

(2): Since X is strictly $(P, \mathcal{B}_1)$-perfect, then the function $\alpha(\omega_1, \omega_2) = (X(\omega_1), X(\omega_2))$ is $(P \otimes P, \mathcal{B}_1 \otimes \mathcal{B}_1)$-perfect with distribution $Q \otimes Q$ and since ξ is strictly $(Q, \mathcal{B}_2)$-perfect, then the function $\beta(x, y) = (\xi(x), \xi(y))$ is $(Q \otimes Q, \mathcal{B}_2 \otimes \mathcal{B}_2)$-perfect. So by (1), we have that $\beta(\alpha(\omega_1, \omega_2))$ is $(P \otimes P, \mathcal{B}_2 \otimes \mathcal{B}_2)$-perfect and since $\beta(\alpha(\omega_1, \omega_2)) = (Y(\omega_1), Y(\omega_2))$, we conclude that Y is strictly $(P, \mathcal{B}_2)$-perfect,

(3): Since $P_Y = Q_\xi$ and Y is $(P, \mathcal{B}_2)$-perfect, then by Lem.A.10.4 we have

$$Q_\xi^*(D) = P_Y^*(D) = P^*(Y^{-1}(D)) = P^*(X^{-1}(\xi^{-1}(D))) \le Q^*(\xi^{-1}(D)) \le Q_\xi^*(D)$$

for all $D \subseteq M_2$. Thus, we see that ξ is $(Q, \mathcal{B}_2)$-perfect.

(4): Let us define $\alpha : \Omega \times \Omega \to M_1 \times M_1$ and $\beta : M_1 \times M_1 \to M_2 \times M_2$ as in the proof of (2). Since $(Y(\omega_1), Y(\omega_2)) = \beta(\alpha(\omega_1, \omega_2))$ and Y is strictly $(P, \mathcal{B}_2)$-perfect, then by (3) we have that $\beta(x, y) = (\xi(x), \xi(y))$ is $(Q \otimes Q, \mathcal{B}_2 \otimes \mathcal{B}_2)$-perfect. Hence, we see that ξ is strictly $(Q, \mathcal{B}_2)$-perfect.

(5): Let $C \subseteq M_1$ be given. Since the restriction of ξ to the $M_1 \setminus N$ is injective, we have $C \setminus N = \xi^{-1}(\xi(C \setminus N)) \setminus N$ and since Y is perfect and $X^{-1}(N)$ is a P-null set, then by Lem.A.10.4 we have

$$\begin{aligned} Q^*(C) &= Q^*(C \setminus N) = Q^*(\xi^{-1}(\xi(C \setminus N))) \le Q_\xi^*(\xi(C \setminus N)) \\ &= P_Y^*(\xi(C \setminus N)) = P^*(Y^{-1}(\xi(C \setminus N)) = P^*(X^{-1}(\xi^{-1}(\xi(C \setminus N)))) \\ &\le P^*(X^{-1}(C \cup N)) = P^*(X^{-1}(C) \cup X^{-1}(N)) = P^*(X \in C) \end{aligned}$$

Hence, by Lem.A.10.4 we see that X is $(P, \mathcal{B}_1)$-perfect.

(6): Let $\alpha : \Omega \times \Omega \to M_1 \times M_1$ and $\beta : M_1 \times M_1 \to M_2 \times M_2$ be defined as in the proof of (2). Since $(Y(\omega_1), Y(\omega_2)) = \beta(\alpha(\omega_1, \omega_2))$ and $\beta(x, y) = (\xi(x), \xi(y))$ is injective on the set $M_1 \times (M_1 \setminus N)$, we see that (6) follows from (5).

(7): Let us define $J(\omega) := \eta(X(\omega))$ for all $\omega \in \Omega$. Since $J(\omega) = \omega$ P-a.s., then by (B.2) we have that $J : \Omega \to (\Omega, \mathcal{F})$ is strictly $(P, \mathcal{F})$-perfect. So by (6) we conclude that X is strictly $(P, \mathcal{B}_1)$-perfect.

(8): Immediate consequence of (7). $\square$

Lemma B.7: *Let $(M_i, \mathcal{B}_i, Q_i)$ be a probability space for all i in the non-empty set I . Let $\alpha \subseteq I$ be a given non-empty set, let $(M^\alpha, \mathcal{B}^\alpha, Q^\alpha) := (\prod_{i \in \alpha} M_i, \otimes_{i \in \alpha} \mathcal{B}_i, \otimes_{i \in \alpha} Q_i)$ denote the product space and let $\pi_\alpha(x) := (x_i)_{i \in \alpha}$ for $x = (x_i)_{i \in I} \in M^I$ denote the projection of M^I onto M^α . Then the projection $\pi_\alpha : (M^I, \mathcal{B}^I) \to (M^\alpha, \mathcal{B}^\alpha)$ is measurable with Q^I-distribution Q^α and we have.*

(1) $\pi_\alpha : M^I \to (M^\alpha, \mathcal{B}^\alpha)$ *is strictly $(Q^I, \mathcal{B}^\alpha)$-perfect and there exists a useful regular probability of Q^I given π_α*

Let $(M,\mathcal{B})$ *be a measurable space and let* $X:\Omega\to M$ *be random function. Then we have*

(2) *If* X *is discrete, then* X *is strictly* $(P,\mathcal{B})$*-perfect*

(3) *If* $\mathcal{B}$ *is a* σ*-discrete* σ*-algebra, then* X *is* $(P,\mathcal{B})$*-perfect if and only if for every non-empty* $\mathcal{B}$*-atom* $D\subseteq M$ *we have* $P^*(X=u)=P_X(D)$ *for all* $u\in D$ *and if so, then we have that* X *is strictly* $(P,\mathcal{B})$*-perfect*

Proof: By the definitions of the product σ-algebra and the product measure, we have that π_α is $(\mathcal{B}^I,\mathcal{B}^\alpha)$-measurable with Q^I-distribution Q^α .

(1): By Thm.B.5.4, it suffices to show that there exists a useful regular conditional distribution of Q^I given π_α . If $\alpha=I$, then $\pi_\alpha=\pi_I$ is the identity map on M^I and $Q(B\mid x):=1_B(x)$ is a useful regular conditional probability of Q^I given π_I . So suppose that $\alpha\neq I$ and let us define $\beta=I\setminus\alpha$. Then we have $M^I=M^\alpha\times M^\beta$ and $Q^I=Q^\alpha\otimes Q^\beta$. Let $F_x:=\{y\in M^\beta\mid(x,y)\in F\}$ denote the x-section of F whenever $F\subseteq M^I=M^\alpha\times M^\beta$ and $x\in M^\alpha$. By Fubini's theorem we have that $Q(F\mid x):=Q^\beta(F_x)$ is a regular conditional distribution of Q^I given π_α . Let $x\in M^\alpha$ and $F\in\mathcal{B}^I$ be given such that $\pi_\alpha^{-1}(x)\subseteq F$. Since $\pi_\alpha^{-1}(x)=\{x\}\times M^\beta$, we have $F_x=M^\beta$ and so we see that $Q(F\mid x)=1$. Hence, we have $Q^*(\pi_\alpha^{-1}(x)\mid x)=1$ for all $x\in M^\alpha$; that is, $Q(F\mid x)$ is a useful regular conditional probability of Q^I given π_α which completes the proof of (1).

(2): Since X is discrete, there exists a countable set $D\subseteq M$ such that $P_X(D)=1$ and $\{q\}\in\mathcal{B}$ and $P(X=q)>0$ for all $q\in D$. If we define $Q(F\mid u):=P(F)$ for all $(F,u)\in\mathcal{F}\times(M\setminus D)$ and

$$Q(F\mid u):=\sum_{q\in D}P(F\mid X=q)\,1_{\{q\}}(u)\quad\forall\,(F,u)\in\mathcal{F}\times D$$

then the reader easily verifies that Q is a useful, regular conditional probability of P given X . So by Thm.B.5.4 we conclude that X is strictly $(P,\mathcal{B})$-perfect.

(3): Let $(D_i\mid i\in I)$ be a disjoint, countable, partition of M such that $\mathcal{B}=\sigma(D_i\mid i\in I)$ and $D_i\neq\emptyset$ for all $i\in I$. Then the sets (D_i) are the non-empty atoms of $\mathcal{B}$. Suppose that X is $(P,\mathcal{B})$-perfect. Let $i\in I$ and $u\subseteq D_i$ be given. Since D_i is a $\mathcal{B}$-atom, we have $P_X^*(\{u\})=P_X(D_i)$ and since X is perfect, we have $P^*(X=u)=P_X^*(\{u\})=P_X(D_i)$. Conversely, suppose that $P^*(X=u)=P_X(D_i)$ for all $i\in I$ and all $u\in D_i$ and let us define

$$Q(F\mid u):=\sum_{i\in I}P(F\mid X\in D_i)\,1_{D_i}(u)\quad\forall\,F\in\mathcal{F}\,\forall\,u\in M$$

where $P(F\mid X\in D_i)$ is the ordinary conditional probability if $P(X\in D_i)>0$ and $P(F\mid X\in D_i):=P(F)$ if $P(X\in D_i)=0$. Then the reader easily verifies that Q is a regular probability of P given X . Let J denote the set of all $i\in I$ satisfying $P(X\in D_i)>0$ and let us define $M_0:=\cup_{i\in J}D_i$. Then we have $P_X(M_0)=1$. Let $u\in M_0$ and $F\in\mathcal{F}$ be given such that

$\{X = u\} \subseteq F$. Then there exists a unique $i \in J$ such that $u \in D_i$ and since $P^*(X = u) = P_X(D_i)$ and $F \cap \{X \in D_i\}$ is a P-measurable set containing $\{X = u\}$, we have $P(F \cap \{X \in D_i\}) = P_X(D_i) = P(X \in D_i)$. So by the definition of Q , we have $Q(F \mid u) = 1$ for all $u \in M_0$ and all $F \in \mathcal{F}$ satisfying $F \supseteq \{X = u\}$. Hence, we have $Q^*(X^{-1}(u) \mid u) = 1$ for all $u \in M_0$ and since $P_X(M_0) = 1$, we conclude that Q is a useful, regular conditional probability of P given X . So by Thm.B.5.4 we conclude that X is strictly $(P, \mathcal{B})$-perfect. $\square$

References:

[1] Andersen, N.T. (1985): *The calculus of non-measurable functions and sets*, Var. Publ. Ser. No.36, Institute of Mathematics, University of Aarhus

[2] Bockstein, M. (1948): *Un théoreme de séparabilité pour les produits topologiques*; Fund. Math. 35, p.242–246

[3] Dudley, R.M. (1984): *A course on empirical processes*, LNS 1097 p.2–141, Springer Verlag, Berlin, New York and London

[4] Dudley, R.M. (1999): *Uniform central limits theorems*, Cambridge University Press, Cambridge and New York

[5] Hoffmann-Jørgensen, J. (1987): *The general marginal problem*, in *Functional Analysis II* (edited by S. Kurepa, H. Kraljević & D. Butković), LNS 1242 p.77–367, Springer Verlag, Berlin, New York and London

[6] Hoffmann-Jørgensen, J. (1993): *Maximal inequalities for random elements*, in *Functional Analysis IV* (edited by S. Kurepa, H. Kraljević & D. Butković), Var. Publ. Ser. No.43 p.105–158, Institute of Mathematics, University of Aarhus

[7] Hoffmann-Jørgensen, J. (1994): *Probability with a view toward statistics*, Vol.I, Chapman & Hall, New York and London

[8] Hoffmann-Jørgensen, J. (1994): *Probability with a view toward statistics*, Vol.II, Chapman & Hall, New York and London

[9] Hoffmann-Jørgensen, J. (1999): *The Glivenko-Cantelli theorem and the Ranga Rao theorem*, in *Functional Analysis VI* (edited by S. Kurepa, H. Kraljević & D. Butković), Var. Publ. Ser. No.43 p.65–290, Institute of Mathematics, University of Aarhus

[10] Kellerer, H.G. (1968): *Stetige Funktionen auf Produkträumen*; Arch. Math. 19, p.79–82

[11] Ledoux, M. & Talagrand, M. (1991): *Probability in Banach spaces*, Springer Verlag, Berlin, New York and London

[12] Pollard, D. (1984): *Convergence of stochastic processes*, Springer Verlag, Berlin, New York and London

[13] Pollard, D. (1990): *Empirical processes: Theory and application*, NSF-CBMS Regional Conference Ser. in Prob. and Stat., Vol.2, Institute of Mathematical Statistics, Hayward California

[14] Rogers, C.A. et al. (1980): *Analytic sets*, Academic Press, London and New York

[15] van der Vaart & A.W & Wellner, J.A. (2000): *Weak convergence and empirical processes*, Springer Verlag, Berlin, New York and London (2nd printing)