Complete Lattices of Probability Measures with Applications to Martingale Theory

Robert P. Kertz*
School of Mathematics, Georgia Institute of Technology
Atlanta, Georgia 30332, U.S.A.
and

Uwe Rösler**

Mathematisches Seminars, Christian-Albrechts-Universität Kiel
Ludewig-Meyn-Str.4, D-24098 Kiel, German

Abstract

The set of probability measures on $\mathbb{R}$ with the stochastic order and the set of right-tail integrable probability measures on $\mathbb{R}$ with the convex order form complete lattices. Connections of these lattice structures to martingale theory and to the Hardy-Littlewood maximal function are exhibited.

AMS 1991 Classifications: Primary 60E15, 60G44; Secondary 60G40, 28A33

Key words and phrases: Martingales, stochastic order, convex order, complete lattices of probability measures, suprema of sets of probability measures, Hardy and Littlewood maximal function.

*Supported in part by NSF grants DMS-88-01818 and DMS-92-09586.
**This research was initiated while this author was at Institut für Mathematische Stochastik, Georg-August Universität, Göttingen.

1 Introduction

The comparison of random variables (r.v.'s) or distributions of random variables leads in a natural way to orderings of probability measures (p.m.'s). For probability measures μ_1 and μ_2 on $\mathbb{R}$, μ_1 is said to be smaller than μ_2

in a given order, if $\int f d\mu_1 \leq \int f d\mu_2$ for all functions f in a given associated class of real-valued functions on $\mathbb{R}$, provided both integrals exist. For the stochastic order the class of increasing functions is used, for the c-convex order increasing convex functions, and for the k-convex order convex functions.

This paper identifies lattice structures of the stochastic and convex partial orders, interprets these structures in terms of martingales and submartingales, and provides a new connection between these structures and Hardy-Littlewood functions. Completeness of the lattice structures is exhibited, and is applied to ensure existence and develop properties of infima and suprema of sets of p.m.'s connected to martingale theory.

These definitions for ordering probability measures allow the comparison of certain features of stochastic processes. Examples are the waiting or service time for queueing systems (see e.g. Stoyan [20] and references therein), the lifetime or reliability of components in complex systems (see e.g. Marshall and Proschan [17]), and rates of return on stocks or mutual funds from a given group (see e.g. Levy [16]). In statistics the stochastic order appears already in the Neyman-Pearson theorem; the most powerful test is the largest r.v. in the stochastic order within the set of tests to a given level (Lehman [15]). The convex order (also called the Bishop-de Leeuw order) is crucial for Choquet theory (Choquet [2], Phelps [19], Meyer [18]). In potential theory, these orders appear in descriptions of balayage, dilation and fusion of measures (Meyer [18], Elton and Hill [6]). A connection to martingale theory was made by Strassen [21]; see Section 5 of this paper. It is useful to consult a reference such as Durrett's book [5] for the context in which additional connections of Hardy-Littlewood functions to martingale theory can be seen. Other examples are scattered throughout the literature.

Our own interest in these orderings arose from 'prophet-gambler' inequalities (Krengel and Sucheston [14], Kertz and Rösler [12]). For a given sequence of games (a given stochastic process), the prophet knows all the outcomes — past, present and future, picks an outcome at any particular instant of time, and receives a reward with a monotone-increasing dependence upon the outcome for that game. The outcome chosen by the prophet is the maximum of the stochastic process, a random variable with associated p.m. ν. The gambler sees the outcomes of the games sequentially, and makes a decision after every game whether to stop or to go on playing; he bases his decision at each instant only on knowledge of the distributions of all the games and the outcomes of this and the past games, and not on knowledge of future outcomes. If he stops, he receives a reward with the same monotone-increasing dependence upon the outcome at the moment he quits. His decision rule is described mathematically by a stopping rule τ, and the outcome upon stopping is the value of the stochastic process at the stopping time; this outcome is a r.v. with associated p.m. μ. Under these restrictions on play, the prophet and the gambler both try to walk away with as much as possible, *on the average.* Of course, in an average-based comparison, the gambler always receives less than the prophet. In the literature, prophet-gambler inequalities seek to identify how close the gambler can be to the prophet over all games of a given type. For example, the inequality might state that the gambler always receives at least a certain proportion of what the prophet receives, e.g. for game outcomes which are independent positive r.v.'s and rewards equal to the outcomes, the proportionality constant is $\frac{1}{2}$ (see Krengel and Sucheston [14]).

A most-interesting collection of sequences of games is the collection of martingales. In a very exact sense, made precise by Hill and Kertz [10], this collection provides the prophet with his greatest uniform advantage. For martingales and convex increasing reward structure, the gambler can use any stopping rule, by the optional sampling theorem. We assume that there is a last outcome and the gambler takes this value. In this context theorems by Hardy and Littlewood and others tell us that the prophet's outcome p.m. ν is stochastically smaller than the Hardy-Littlewood maximal p.m. μ^* (see Hardy and Littlewood [9], Blackwell and Dubins [1], and Dubins and Gilat [4]). There is at least one martingale where the prophet actually chooses outcomes with associated p.m. μ^*, and the gambler chooses outcomes with p.m. only μ. Kertz and Rösler [12], showed that for any ν stochastically between μ and μ^*, there exists a martingale with outcome p.m. ν for the prophet and μ for the gambler.

In main results of this paper given in Section 3, complete lattice structures are given for sets of p.m.'s with the stochastic and convex orders. Specifically, let $\mathcal{P}$ be the set of probability measures on $\mathbb{R}$, $\mathcal{P}_+$ be the subset of p.m.'s with finite first moment integral over $\mathbb{R}^+$, and $\mathcal{P}_r$ be the set of p.m.'s on $\mathbb{R}$ with first moment integral equal to the real number r. It is shown in Section 3 that $(\mathcal{P}, \prec_s)$, $(\mathcal{P}_+, \prec_c)$ and $(\mathcal{P}_r, \prec_k)$ are complete lattices. By complete, we mean that any set bounded below has an infimum and any set bounded above has a supremum. The complete lattice structure for the last two sets appears to be new.

This lattice completeness is applied in Theorems 3.5 and 3.7 to obtain results on infima and suprema of specific sets of p.m.'s in the stochastic and convex orders. These sets of p.m.'s are related to sets of p.m.'s used in the 'prophet vs. gambler' comparisons in Kertz and Rösler [13].

In Section 4 we use the result that two right-tail integrable p.m.'s are c-convex ordered if and only if the associated Hardy-Littlewood maximal p.m.'s are stochastically ordered to obtain a complete lattice structure on the set of Hardy-Littlewood maximal p.m.'s. Through a natural isomorphic representation between this set and a set of concave functions, this natural lattice structure on the set of Hardy-Littlewood maximal p.m.'s is shown to be related to the lattice structure based on the stochastic order.

Martingale connections to these orders on p.m.'s and to the suprema and infima results are given in Section 5. In particular, in Theorem 5.2 an extension of a representation result for submartingales given by Strassen [21] is given in a generalized-submartingale context. Martingale interpretations of Theorems 3.5 and 4.9 are given in Theorems 5.3 and 5.5.

2 The stochastic and convex partial orders

On the set $\mathcal{P}(\mathbb{R})$ of all probability measures on $\mathbb{R}$ we introduce the three relations $\prec_s, \prec_c$ and $\prec_k$. Define $\mu_1 \prec_s \mu_2$, $\mu_1 \prec_c \mu_2$, $\mu_1 \prec_k \mu_2$ iff

$$\int f(x)d\mu_1(x) \leq \int f(x)d\mu_2(x) \tag{2.1}$$

for all functions $f : \mathbb{R} \to \mathbb{R}$ for which the integrals are well defined and f is increasing, respectively increasing and convex, respectively convex. We use increasing for $x < y \Rightarrow f(x) \leq f(y)$ and positive for $x \geq 0$. If we need the strict inequality, we use strictly increasing or strictly positive.

The relation $\prec_s$ is called the stochastic order; and the relation $\prec_k$ is called the convex order (see Stoyan [20] for an account). We shall use c-convex and k-convex to distinguish the two orders $\prec_c$ and $\prec_k$. These relations are obviously transitive and reflexive. If two p.m.'s are ordered in the stochastic order or the k-convex order, then they are ordered in the c-convex order.

The set of positive functions $f : \mathbb{R} \to \mathbb{R}$ which are increasing, respectively increasing and convex, respectively convex, form a cone. The extremal elements are, in addition to the constant function 1, the step functions in $x \in \mathbb{R}$, $1_{t \leq x}$, $1_{t<x}$, $t \in \mathbb{R}$; respectively $(x-t)^+$, $t \in \mathbb{R}$; respectively $(x-t)^+$, $(t-x)^+$, $t \in \mathbb{R}$. The $+$ denotes the positive part.

The integral over the step functions is always finite. Therefore, the stochastic order is in fact a partial order. We consider the relation $\prec_s$ on the set $\mathcal{P}(\mathbb{R})$. For the $\prec_c$ relation we require that the p.m.'s under comparison have a finite integral with at least one non constant, convex and increasing function f. But then for such a p.m., the integral for any extremal function $(x-t)^+$ is finite. This natural requirement leads to the smaller class

$$\mathcal{P}_+ := \{\mu \in \mathcal{P}(\mathbb{R}) : \int_0^\infty x d\mu(x) < \infty\}$$

of probability measures. The relation $\prec_c$ is in fact a partial order on this class. Obviously the relation $\prec_k$ is even stronger than the relation $\prec_c$. Therefore, we require that p.m.'s under this comparison have finite integral with at least one convex function f which does not belong to the previous class.

But then for such a p.m., the integral for any of the extremal functions is finite. The natural class of probability measures here is

$$\mathcal{P}_f = \{\mu \in \mathcal{P}(\mathbb{R}) : \int |x| d\mu(x) < \infty\}.$$

The relation $\prec_k$ is a partial order on this set.

For each p.m. μ on $\mathbb{R}$, $F = F_\mu$ denotes the distribution function of μ. We always drop indices like μ whenever possible. The left continuous inverse $F^{-1} : (0,1) \to \mathbb{R}$ is as usual defined by $F^{-1}(w) = \inf\{z : \; F(z) \geq w\}$. If appropriate we use the continuous extension $F^{-1} : [0,1] \to \mathbb{R} \cup \{-\infty\} \cup \{+\infty\}$.

The following basic results are well known (see e.g. Stoyan [20])

Lemma 2.1 *Each of the following is equivalent for $\mu_1, \mu_2 \in \mathcal{P}(\mathbb{R})$.*

(i) $\mu_1 \prec_s \mu_2$.

(ii) $\int f d\mu_1 \leq \int f d\mu_2$ *for all positive increasing functions* $f : \mathbb{R} \to \mathbb{R}$.

(iii) $\int f d\mu_1 \leq \int f d\mu_2$ *for all (extremal) indicator functions* $f = 1_{t \leq .}$, $t \in \mathbb{R}$.

(iv) $F_{\mu_2} \leq F_{\mu_1}$ *pointwise.*

(v) $F^{-1}_{\mu_1} \leq F^{-1}_{\mu_2}$ *pointwise.*

A similar result holds for the $\prec_c$ order (see again Stoyan [20]).

Lemma 2.2 *Each of the following is equivalent for $\mu_1, \mu_2 \in \mathcal{P}_+$.*

(i) $\mu_1 \prec_c \mu_2$.

(ii) $\int f d\mu_1 \leq \int f d\mu_2$ *for all positive increasing convex functions* $f : \mathbb{R} \to \mathbb{R}$.

(iii) $\int f d\mu_1 \leq \int f d\mu_2$ *for all (extremal) functions* $f(.) = (. - t)^+$, $t \in \mathbb{R}$.

(iv) $\int_.^\infty (1 - F_{\mu_1}(s))ds \leq \int_.^\infty (1 - F_{\mu_2}(s))ds$ *pointwise.*

(v) $\int_.^1 F^{-1}_{\mu_1}(u)du \leq \int_.^1 F^{-1}_{\mu_2}(u)du$ *pointwise.*

And a similar result holds for the $\prec_k$ order.

Lemma 2.3 *Each of the following is equivalent for $\mu_1, \mu_2 \in \mathcal{P}_f$.*

(i) $\mu_1 \prec_k \mu_2$.

(ii) $\int f d\mu_1 \leq \int f d\mu_2$ *for all positive convex functions* $f : \mathbb{R} \to \mathbb{R}$.

(*iii*) $\int f d\mu_1 \leq \int f d\mu_2$ *for all (extremal) functions* $f(.) = (. - t)^+$, *or* $f(.) = (t - .)^+$, $t \in I\!R$.

(*iv*) $\int_.^\infty (1 - F_{\mu_1}(s))ds \leq \int_.^\infty (1 - F_{\mu_2}(s))ds$ *pointwise and* $\int x dF_{\mu_1}(x) = \int x dF_{\mu_2}(x)$.

(*v*) $\int_.^1 F_{\mu_1}^{-1}(u)du \leq \int_.^1 F_{\mu_2}^{-1}(u)du$ *pointwise and* $\int_0^1 F_{\mu_1}^{-1}(u)du = \int_0^1 F_{\mu_2}^{-1}(u)du$.

(*vi*) $\mu_1 \prec_c \mu_2$ *and* $\int x d\mu_1(x) = \int x d\mu_2(x)$.

For the equivalence (v) in each of Lemmas 2.1 and 2.2, see Lemma 1.8 of Kertz and Rösler [13]. Lemma 2.3 shows $\mu_1 \prec_k \mu_2$ for $\mu_1, \mu_2 \in \mathcal{P}_f$ already implies that μ_1 and μ_2 have the same finite expectation. Therefore it is reasonable to consider the partial order $\prec_k$ only on a set

$$\mathcal{P}_r := \{\mu \in \mathcal{P}(I\!R) : \int x d\mu(x) = r\}$$

for fixed $r \in I\!R$. But on $\mathcal{P}_r$ the partial orders $\prec_c$ and $\prec_k$ are the same. An easy criterion for k-convexity is the cut criterion by Karlin and Novikoff [11].

Proposition 2.4 *Let* μ_1, μ_2 *be in* $\mathcal{P}_r$ *and the function* $F_{\mu_1} - F_{\mu_2}$ *be negative from minus infinity to some point and then positive. Then* $\mu_1 \prec_k \mu_2$.

To complete the picture, and for later use, we define the relation $\prec_d$ on $\mathcal{P}(I\!R)$ by $\mu_1 \prec_d \mu_2$ iff (2.1) holds for all functions $f : I\!R \to I\!R$ for which the integrals are well defined and f is decreasing and convex. The relation $\prec_d$ is a partial order on the set of p.m.'s

$$\mathcal{P}_- := \{\mu \in \mathcal{P}(I\!R) : \int_{-\infty}^0 x d\mu(x) > -\infty\}.$$

We use the terminology 'd–convex order' when referring to the $\prec_d$ order. The following gives characterizations of this $\prec_d$ order, and an immediate connection to the $\prec_c$ and $\prec_k$ orders. The proof is immediate from the definitions and Lemmas 2.2 and 2.3.

Lemma 2.5 *(a) Each of the following is equivalent for* $\mu_1, \mu_2 \in \mathcal{P}_-$.

(*i*) $\mu_1 \prec_d \mu_2$.

(*ii*) $\int f d\mu_1 \leq \int f d\mu_2$ *for all positive decreasing convex functions* $f : I\!R \to I\!R$.

(*iii*) $\int f d\mu_1 \leq \int f d\mu_2$ *for all functions* $f(.) = (. - t)^- = (t - .)^+$, $t \in I\!R$.

(*iv*) $\int_{-\infty}^. F_{\mu_1}(s)ds \leq \int_{-\infty}^. F_{\mu_2}(s)ds$ *pointwise.*

(v) $\int_0^{\cdot} F_{\mu_2}^{-1}(u)du \leq \int_0^{\cdot} F_{\mu_1}^{-1}(u)du$ *pointwise.*

(vi) $\mu_{-X_1} \prec_c \mu_{-X_2}$, *where* X_1 *and* X_2 *are r.v.'s with associated p.m.'s* μ_1 *and* μ_2 *respectively.*

(b) For $\mu_1, \mu_2 \in \mathcal{P}_f$, $\mu_1 \prec_k \mu_2$ *if and only if both* $\mu_1 \prec_c \mu_2$ *and* $\mu_1 \prec_d \mu_2$.

3 The lattice structures

We shall first discuss $(\mathcal{P}(I\!R), \prec_s)$, $(\mathcal{P}_+, \prec_c)$ and $(\mathcal{P}_f, \prec_k)$ as lattices. We call a lattice *complete*, if (i) any nonempty subset which is bounded above has a supremum (least upper bound) and (ii) any nonempty subset which is bounded below has an infimum (greatest lower bound) (for reference, see e.g. Chapter 6 of Gleason [8]). As shown in Proposition 6-5.1 of [8], if either of (i) or (ii) in this definition holds, then the other also holds. In the following the symbols $\vee$ and $\wedge$ denote the supremum and infimum of real numbers in the usual order.

Lemma 3.1 $(\mathcal{P}(I\!R), \prec_s)$ *is a complete lattice. The operations of supremum* $\vee_s$ *and infimum* $\wedge_s$ *are given for* $\mu_1, \mu_2 \in \mathcal{P}(I\!R)$ *by*

$$(\mu_1 \vee_s \mu_2)([x, \infty)) = \mu_1([x, \infty)) \vee \mu_2([x, \infty))$$

$$(\mu_1 \wedge_s \mu_2)([x, \infty)) = \mu_1([x, \infty)) \wedge \mu_2([x, \infty))$$

for all $x \in I\!R$.

Proof: It is straightforward to show that $(\mathcal{P}(I\!R), \prec_s)$ is a lattice. To show this lattice is complete, let K be any nonempty subset of $\mathcal{P}(I\!R)$ which is bounded from above; then the supremum of K, $\vee_s\{\mu : \mu \in K\}$, is the left-continuous modification of $\vee_s\{\mu : \mu \in K\}([x, \infty)) = \vee\{\mu[x, \infty) : \mu \in K\}$. Similarly, if K is any nonempty subset of $\mathcal{P}(I\!R)$ which is bounded from below, then the infimum of K, $\wedge_s\{\mu : \mu \in K\}$ is $\wedge_s\{\mu : \mu \in K\}([x, \infty)) = \wedge\{\mu[x, \infty) : \mu \in K\}$. q.e.d.

The space $(\mathcal{P}(I\!R), \prec_s)$ has no smallest or largest element. However, if we extend $I\!R$ to $\overline{I\!R}$ by adding plus and minus infinity with the one point compactification, then $(\mathcal{P}(\overline{I\!R}), \prec_s)$ has the point measures on minus infinity and plus infinity as smallest and largest elements.

We could find no reference that $(\mathcal{P}_+, \prec_c)$ is a complete lattice, and shall therefore prove that result. An earlier version of the first part of this result was given in [13]; to provide a simpler argument, and for the reader's convenience, we supply the entire argument.

Let $\mathcal{C}$ be the space of all functions $\psi : I\!R \to I\!R$ satisfying (i) ψ is convex and increasing; (ii) $\lim_{t \uparrow \infty}(\psi(t) - t) = 0$; and (iii) $\lim_{t \downarrow -\infty}(\psi(t+h) - \psi(t)) = 0$ for all $h \in I\!R$. The usual pointwise order on functions is taken to be the partial order on $\mathcal{C}$, denoted by $\prec$. We first identify the lattice structure on $(\mathcal{C}, \prec)$, and then exhibit a lattice isomorphism between $(\mathcal{C}, \prec)$ and $(\mathcal{P}_+, \prec_c)$. Recall (e.g. Choquet [2]) that the *lower convex envelope* of a set S of real valued functions, denoted $conv(S)$, is the lower convex envelope of the pointwise infimum of the functions in S. It is the greatest convex function pointwise smaller than or equal to any function in S. Define the supremum $\sqcup$ and the infimum $\sqcap$ on $\mathcal{C}$ by

$$(\psi_1 \sqcup \psi_2)(t) = \psi_1(t) \vee \psi_2(t) \quad \text{and} \quad (\psi_1 \sqcap \psi_2)(t) = (conv(\{\psi_1, \psi_2\}))(t).$$

for all $t \in I\!R$.

Lemma 3.2 *The space $(\mathcal{C}, \prec)$ is a complete lattice. The supremum and infimum are given by the operations $\sqcup, \sqcap : \mathcal{C}\text{x}\mathcal{C} \to \mathcal{C}$.*

Proof: The relation $\prec$ is reflexive, anti-symmetric and transitive, and therefore is a partial order. The functions $\sqcup$ and $\sqcap$ are well defined. For the supremum use the pointwise maximum of two convex functions is again convex, and then check the conditions for functions to be in $\mathcal{C}$. We show that if $\psi_1,\ \psi_2 \in \mathcal{C}$, then $\psi_1 \sqcap \psi_2 \in \mathcal{C}$. Recall any function has a unique greatest lower convex envelope. Therefore the function $\psi_1 \sqcap \psi_2$ is well defined and convex, increasing. From the defining properties (i) and (ii) of $\mathcal{C}$, we obtain $t \leq \psi(t)$ for all $t \in I\!R$, $\psi \in \mathcal{C}$. An immediate consequence is $\lim_{t\to\infty}((\psi_1 \sqcap \psi_2)(t) - t) = 0$. For the last defining property (iii) of $\mathcal{C}$, observe that

$$\begin{aligned} 0 &\leq (\psi_1 \sqcap \psi_2)(t+h) - (\psi_1 \sqcap \psi_2)(t) \\ &\leq |\psi_1(t+h) - \psi_1(t)| + |\psi_2(t+h) - \psi_2(t)| \ \to_{t\to-\infty}\ 0. \end{aligned}$$

By construction $\sqcup$ and $\sqcap$ are in fact the least upper bound and greatest lower bound as required. This proves the lattice structure.

Let K be any nonempty subset of $\mathcal{C}$ which is bounded from above by some $g_1 \in \mathcal{C}$. Then the pointwise supremum f of all functions in K is a well defined, convex and increasing function. Further

$$0 \leq \limsup_{t\to\infty}(f(t) - t) \leq \limsup_{t\to\infty}(g_1(t) - t) = 0.$$

The function $f(t+h) - f(t)$ is increasing in $t \in I\!R$, for each $h \in I\!R^+$ fixed. The limit as $t \to -\infty$ exists and is some positive value, called v. It is easy to show that $v > 0$ would imply that each k in K is not real valued (this uses that $\lim_{t\downarrow-\infty}(k(t+h) - k(t)) = 0$ for all $\ell \in I\!R$); so we must have $v = 0$. It follows easily that $\lim_{t\to-\infty}(f(t+h) - f(t)) = 0$ for *all* $h \in I\!R$. Thus $f \in \mathcal{C}$; it is immediate that f is the supremum (least upper bound) of K in $\mathcal{C}$.

Now let K be any nonempty subset of $\mathcal{C}$ that is bounded from below by some $g_2 \in \mathcal{C}$. The pointwise infimum of all functions in K is well defined. Call the lower convex envelope f. Then f is convex and increasing. Further

$$0 \leq \limsup_{t\to\infty}(f(t)-t) \leq \limsup_{t\to\infty}(k(t)-t) = 0$$

for any $k \in K$. The function $f(t+h)-f(t)$ is increasing in $t \in \mathbb{R}$, for each $h \in \mathbb{R}^+$ fixed. The limit as t tends to minus infinity exists and is some positive value, called v. As before, one may show that if $v > 0$, then g_2 would not be real valued; so we must have $v = 0$. It again follows easily that $\lim_{t\to-\infty}(f(t+h)-f(t)) = 0$ for *all* $h \in \mathbb{R}$. Therefore f is the greatest lower bound of K in $\mathcal{C}$. q.e.d.

Let ψ be in $\mathcal{C}$. Denote its right-sided derivative by ψ' (or one may work with its Radon-Nikodym derivative). Then ψ' is (or can be chosen to be) an increasing right continuous function. We may think of ψ' as a distribution function. Indeed, by $\psi \in \mathcal{C}$ and the representation $\psi(t)-\psi(0) = \int_0^t \psi'(s)ds$, conclude that

$$\lim_{s\to-\infty}\psi'(s) = 0, \quad \lim_{s\to\infty}\psi'(s) = 1, \quad 0 \leq \int_0^\infty (1-\psi'(s))ds < \infty;$$

and we also obtain the representation $\psi(t) - t = \int_t^\infty (1-\psi'(s))ds$. Denote by Γ the map from ψ' to ψ. If we identify ψ' and its associated p.m., it is clear that Γ is a map from $\mathcal{P}_+$ onto $\mathcal{C}$.

Theorem 3.3 *The map Γ is a lattice isomorphism from $(\mathcal{P}_+, \prec_c)$ to $(\mathcal{C}, \prec)$.*

Proof: The map $\Gamma : \mathcal{P}_+ \to \mathcal{C}$ is well defined. It is easy to show Γ is a one-to-one mapping. By use of the equivalences in Lemma 2.2, it is also straightforward to show that Γ is order-preserving, i.e., $\mu_1 \prec_c \mu_2 \iff \Gamma(\mu_1) \prec \Gamma(\mu_2)$. q.e.d.

The result that the mapping Γ defines a one-to-one mapping from $\mathcal{P}_f$ to a subset of $\mathcal{C}$ was proved by Gilat [7]. As a consequence of Lemma 3.2 and Theorem 3.3, we have the following.

Theorem 3.4 *$(\mathcal{P}_+, \prec_c)$ is a complete lattice. The supremum and infimum operations $\vee_c$ and $\wedge_c$ are given for $\mu_1, \mu_2 \in \mathcal{P}_+$ by*

$$\mu_1 \vee_c \mu_2 = \Gamma^{-1}(\Gamma(\mu_1) \sqcup \Gamma(\mu_2)) \quad \textit{and} \quad \mu_1 \wedge_c \mu_2 = \Gamma^{-1}(\Gamma(\mu_1) \sqcap \Gamma(\mu_2)).$$

The lattices $(\mathcal{C}, \prec)$ and $(\mathcal{P}_+, \prec_c)$ have no smallest or largest element. If we compactify $\mathbb{R}$ as before, the point measures at minus infinity and at plus infinity are respectively the smallest and greatest elements.

The space $(\mathcal{P}_f, \prec_k)$ is not a lattice. The space $(\mathcal{P}_r, \prec_k) = (\mathcal{P}_r, \prec_c)$ is a lattice. Notice this lattice is complete and has a smallest element, the point measure at r, but no greatest element.

The next theorems are interesting consequences of the completeness property of these lattices. These results have connections to martingale theory which are given in Section 5.

For these theorems, we introduce the notation

$$M_{ab} := \{\mu \in \mathcal{P}(\mathbb{R}) : \mu_1 \prec_a \mu \prec_b \mu_2\}$$

where $\prec_a$ denotes either $\prec_s$, the stochastic order, with $\mu_1 \in \mathcal{P}(\mathbb{R})$; $\prec_c$, the c–convex order, with $\mu_1 \in \mathcal{P}_+$; $\prec_d$, the d–convex order, with $\mu_1 \in \mathcal{P}_-$; or $\prec_k$, the k–convex order, with $\mu_1 \in \mathcal{P}_r$. Similarly $\prec_b$ denotes either $\prec_s$, $\prec_c$, $\prec_d$ or $\prec_k$, with μ_2 in $\mathcal{P}(\mathbb{R})$, $\mathcal{P}_+, \mathcal{P}_-$ or $\mathcal{P}_r$ respectively.

Theorem 3.5 *Assume μ_1 and μ_2 are p.m.'s in $\mathcal{P}_+$ which satisfy $\mu_1 \prec_c \mu_2$. Then M_{sc} contains a maximal element in the stochastic order. Assume, in addition, that $\mu_2 \in \mathcal{P}_r$. Then the following hold.*

(i) M_{sc} and M_{sk} contain the same maximal elements in the stochastic order.

(ii) M_{sk} and M_{ck} have the same infimum in the k–convex order. This infimum γ, contained in both sets, has associated left continuous inverse function $F_\gamma^{-1}(w) = x_0 \vee F_{\mu_1}^{-1}(w)$ for $w \in (0,1)$, where x_0 is the number in $[-\infty, \infty)$ chosen so that $\int_0^1 x_0 \vee F_{\mu_1}^{-1}(w)dw = \int_0^1 F_{\mu_2}^{-1}(w)dw$.

(iii) The infimum of M_{sk} in the k–convex order is a maximal element of M_{sk} in the stochastic order.

Proof: For the first conclusion, observe that the set M_{sc} is non empty. Let K be a stochastic–order chain in M_{sc}. The set K is bounded from above in the stochastic order. Indeed, for any $\mu \in K$, and all $w \in (0,1)$, $F_\mu^{-1}(w) \leq (1-w)^{-1}\int_w^1 F_\mu^{-1}(u)du \leq (1-w)^{-1}\int_w^1 F_{\mu_2}^{-1}(u)du =: g(w)$; and the p.m. associated with this r.v. g is a stochastic-order upper bound for K. Let μ' be the stochastic–order supremum of K. Then $\mu_1 \prec_s \mu'$, and $\mu' \prec_c \mu_2$ since $\int f d\mu' = \lim_{K \ni \mu \uparrow} \int f d\mu < \infty$ for any increasing, convex and positive function f. We can now apply Zorn's lemma to obtain that M_{sc} has a maximal element in the stochastic order.

Now, assume in addition that $\mu_2 \in \mathcal{P}_r$. For conclusion (i) it is straightforward to obtain that, in the stochastic order, maximal elements for M_{sk} are also maximal elements for M_{sc}. To obtain that, in the stochastic order, maximal elements for M_{sc} are also maximal elements for M_{sk}, it suffices to show that any maximal element for M_{sc} in the stochastic order is also in M_{sk}. Let μ be such a maximal element for M_{sc}.

Suppose that $\int xd\mu(x) < \int xd\mu_2(x) = r$. Let T be chosen so that $\int xd\mu(x) < \int (x \vee T)d\mu(x) < \int xd\mu_2(x)$. Define p.m. $\mu_T(\cdot) = \mu(-\infty, T)\epsilon_T(\cdot) + \mu(\cdot \cap [T, \infty))$. Then $\mu \prec_s \mu_T$ and $\mu \neq \mu_T$. We show that $\mu_T \prec_c \mu_2$ by showing that $\int (x \vee t)d\mu_T(x) \leq \int (x \vee t)d\mu_2(x)$ for all $t \in I\!R$. Since $\mu \prec_c \mu_2$, we have $\int (x \vee t)d\mu_T(x) = \int (x \vee t)d\mu(x) \leq \int (x \vee t)d\mu_2(x)$ for all $t > T$. For $t \leq T$, it follows that

$$\begin{aligned}\int (x \vee t)d\mu_T(x) &= \int_{x>T} xd\mu(x) + T\mu(-\infty, T] = \int (x \vee T)d\mu(x) \\ &< \int xd\mu_2(x) \leq \int (x \vee t)d\mu_2(x).\end{aligned}$$

But this gives a contradiction to the maximality of μ. Thus, we must have that the maximal element μ satisfies $\mu_1 \prec_s \mu \prec_k \mu_2$, and so μ is in M_{sk}.

Next, conclusion (ii) is proved. The p.m. γ, identified in the statement of the conclusion of this Theorem, is the infimum of M_{sk} in the k-convex order. Indeed, γ clearly satisfies $\mu_1 \prec_s \gamma$ and $\int xd\gamma(x) = r = \int xd\mu_2(x)$; and one shows in a straightforward way (e.g., through Lemma 2.2 (iii)) that $\gamma \prec_c \mu_2$, and that $\gamma \prec_c \mu$ for all $\mu \in M_{sk}$. This gives that $\gamma \in M_{sk}$, and that $\gamma = \inf_{\prec_k} M_{sk}$.

Since M_{ck} is bounded below in the k–convex order by the point measure at r, and $(\mathcal{P}_r, \prec_k)$ is complete, it follows that M_{ck} has an infimum in the k–convex order; call this infimum γ_c. Now, $M_{sk} \subseteq M_{ck}$ and $\gamma_c \prec_k \gamma$. We show that $\gamma_c = \gamma$. ¿From the c-convex order requirements and Lemma 2.2, we know that for all $w \in (0,1)$

$$\int_w^1 F_{\mu_1}^{-1}(u)du \leq \int_w^1 F_{\gamma_c}^{-1}(u)du \leq \int_w^1 F_{\gamma}^{-1}(u)du. \tag{3.1}$$

Let t_0 satisfy $F_{\mu_1}^{-1}(t_0) \leq x_0 < F_{\mu_1}^{-1}(t_0+)$. ¿From the definition of F_γ^{-1} and from (3.1), we obtain that $F_{\gamma_c}^{-1}(w) = F_{\mu_1}^{-1}(w)$ for $w \in (t_0, 1)$. But then from (3.1) we must also have that for all $w \in (0, t_0)$

$$\int_w^{t_0} F_{\gamma_c}^{-1}(u)du \leq \int_w^{t_0} F_{\gamma}^{-1}(u)du = x_0(t_0 - w);$$

and it follows that $F_{\gamma_c}^{-1}(t_0) \leq x_0$. However, γ_c and γ have the same mean; and this implies that $\int_0^{t_0} F_{\gamma_c}^{-1}(u)du = \int_0^{t_0} F_{\gamma}^{-1}(u)du = x_0 t_0$. Thus, $F_{\gamma_c}^{-1}(w) = x_0$ for $w \in (0, t_0)$; and $\gamma_c = \gamma$.

To prove (iii), let $\gamma = \inf_{\prec_k} M_{sk}$, given in part (ii), and suppose there is a p.m. $\bar{\mu}$ in M_{sk} for which $\gamma \prec_s \bar{\mu}$ and $\gamma \neq \bar{\mu}$. Then $F_\gamma^{-1}(w) \leq F_{\bar{\mu}}^{-1}(w)$ for all $w \in (0,1)$; and $F_\gamma^{-1}(w) < F_{\bar{\mu}}^{-1}(w)$ on some set of positive Lebesgue measure. Since both $\bar{\mu}$ and γ are in M_{sk}, this implies that

$$r = \int xd\gamma = \int_0^1 F_\gamma^{-1}(w)dw < \int_0^1 F_{\bar{\mu}}^{-1}(w)dw = \int xd\bar{\mu} = r,$$

a contradiction. It follows that γ is a maximal element of M_{sk} in the stochastic order. q.e.d.

Remark 3.6 Under the assumptions that $\mu_1 \in \mathcal{P}(\mathbb{R})$, $\mu_2 \in \mathcal{P}_r$ and $\mu_1 \prec_c \mu_2$, the sets M_{sk} and M_{ck} satisfy

$$M_{sk} \subseteq M_{ck} = \{\mu \in \mathcal{P}(\mathbb{R}) : \gamma \prec_k \mu \prec_k \mu_2\} \tag{3.2}$$

where γ is the p.m. of Theorem 3.5(ii), the common infimum of sets M_{sk} and M_{ck} in the k–convex order. The containment of (3.2) is clear; and the equality follows from Theorem 3.5. The containment of the two sets of p.m.'s in (3.2) may be strict, as we see in the following example. Let μ_1 and μ_2 be uniformly distributed on the intervals (0,1) and (1/4,5/4) respectively, so that $F_{\mu_1}^{-1}(w) = w$ and $F_{\mu_2}^{-1}(w) = w + (1/4)$ for $w \in (0,1)$. Then $F_{\gamma}^{-1}(w) = x_0 \vee F_{\mu_1}^{-1}(w) = \left(\sqrt{2}/2\right) \vee w$ for $w \in (0,1)$. Now, let μ be the p.m. with left continuous inverse function $F_{\mu}^{-1}(w) = (1/4) + w$ if $c < w \leq 1$, and $= \sqrt{2}/2$ if $0 < w \leq c$, where $c = \left(2\sqrt{2} - 1\right)/2$. Then $\mu \in M_{ck}$; but $\mu \notin M_{sk}$, since it is not true that $\mu_1 \prec_s \mu$.

¿From Lemma 2.5 and Theorems 3.4 and 3.5, and from the fact that $\mu_{X_1} \prec_s \mu_{X_2}$ if and only if $\mu_{-X_2} \prec_s \mu_{-X_1}$, the following results are clear.

Theorem 3.7 *(a)* $(\mathcal{P}_-, \prec_d)$ *is a complete lattice. The supremum and infimum operations* $\vee_d$ *and* $\wedge_d$ *are given for* μ_{X_1}, μ_{X_2} *in* $\mathcal{P}_-$ *by* $\mu_{X_1} \vee_d \mu_{X_2} = \mu_{-X_1} \vee_c \mu_{-X_2}$ *and* $\mu_{X_1} \wedge_d \mu_{X_2} = \mu_{-X_1} \wedge_c \mu_{-X_2}$.

(b) Assume μ_1 *and* μ_2 *be p.m.'s in* $\mathcal{P}_-$ *which satisfy* $\mu_1 \prec_d \mu_2$. *Then* M_{sd} *contains a minimal element in the stochastic order. Assume, in addition,* $\mu_2 \in \mathcal{P}_r$. *Then the following hold.*

(i) M_{sd} *and* M_{sk} *contain the same minimal elements in the stochastic order.*

(ii) M_{sk} *and* M_{dk} *have the same infimum in the k–convex order. This infimum* ρ, *contained in both these sets, has associated left continuous inverse function* $F_{\rho}^{-1}(w) = (-y_0) \wedge F_{\mu_1}^{-1}(w)$ *for* $w \in (0,1)$, *where* y_0 *is the number in* $[-\infty, \infty)$ *chosen so that* $\int_0^1 (-y_0) \wedge F_{\mu_1}^{-1}(w)dw = \int_0^1 F_{\mu_2}^{-1}(w)dw$. *These sets satisfy*

$$M_{sk} \subseteq M_{dk} = \{\mu \in \mathcal{P}(\mathbb{R}) : \rho \prec_k \mu \prec_k \mu_2\},$$

where the inclusion may be strict.

(iii) The infimum of M_{sk} *in the k–conex order is a minimal element of* M_{sk} *in the stochastic order.*

4 The Hardy-Littlewood maximal function

For any p.m. in $\mathcal{P}_+$ with corresponding distribution function F define the *Hardy-Littlewood maximal function* $H^{-1} : (0,1) \to I\!R$ by

$$H^{-1}(w) := (1-w)^{-1} \int_w^1 F^{-1}(u)du.$$

Notice the conditions $\mu \in \mathcal{P}_+$ and $\int_w^1 F_\mu^{-1}(u)du < \infty$ for one (and then for all) $w \in (0,1)$ are equivalent. We use the continuous extension $H^{-1} : [0,1] \to \overline{I\!R}$ if necessary.

Denote the right continuous inverse of H^{-1} by $H : I\!R \to [0,1]$. The function $H = H_\mu$ is a distribution function, continuous except possibly for one jump at the point b, for which $H(b-) \leq 1$, $H(b+) = 1$. H is called the *Hardy-Littlewood maximal distribution function.* We denote the corresponding probability measure by μ^*, and refer to this p.m. as the *Hardy-Littlewood maximal p.m.* associated with μ.

Here we also use $*$ to denote the map from $\mathcal{P}_+$ to the set of probability measures defined by $*(\mu) = \mu^*$. The image of $*$ is the set of all Hardy-Littlewood p.m.'s, denoted by $\mathcal{H}^*$. In this Section we give several isomorphism characterizations of the set $\mathcal{H}^*$. A basic fact is that p.m.'s are related through the c-convex order if and only if their respective Hardy-Littlewood maximal p.m.'s are related through the stochastic order, as stated in the following Lemma.

Lemma 4.1 *For p.m.'s μ_1 and μ_2 in $\mathcal{P}_+$, $\mu_1 \prec_c \mu_2$ is equivalent to $\mu_1^* \prec_s \mu_2^*$.*

Proof: The proof uses equivalences for the stochastic order and for the c–convex order given in Lemmas 2.1 and 2.2. Let F_1, F_2 be the distribution functions of μ_1, μ_2. For simplicity we assume that F_1 and F_2 are strictly increasing and have no jumps; the modifications for general F_1 and F_2 are straightforward. Then for all $t \in I\!R$

$$\begin{aligned}
\mu_1 \prec_c \mu_2 \quad &\Leftrightarrow \quad \int (x-t)^+ F_1(x) \leq \int (x-t)^+ dF_2(x) \\
&\Leftrightarrow \quad \int_{F_1(t)}^1 F_1^{-1}(u)du - t(1-F_1(t)) \leq \int_{F_2(t)}^1 F_2^{-1}(u)du - t(1-F_2(t)) \\
&\Leftrightarrow \quad \int_{F_1(t)}^1 F_1^{-1}(u)du \leq \int_{F_1(t)}^1 F_2^{-1}(u)du + \int_{F_2(t)}^{F_1(t)} (F_2^{-1}(u)-t)du \\
&\Leftrightarrow \quad \int_{F_2(t)}^1 F_1^{-1}(u)du + \int_{F_1(t)}^{F_2(t)} (F_1^{-1}(u)-t)du \leq \int_{F_2(t)}^1 F_2^{-1}(u)du.
\end{aligned}$$

The main point is now to notice $\int_{F_2(t)}^{F_1(t)} (F_2^{-1}(u)-t)du$ and $\int_{F_1(t)}^{F_2(t)} (F_1^{-1}(u)-t)du$ are both always positive. Thus $\mu_1 \prec_c \mu_2 \Leftrightarrow$

$$\Leftrightarrow \quad H_1^{-1}(F_2(t)) + \textit{positive} \leq H_2^{-1}(F_2(t))$$

$$\Leftrightarrow\ H_1^{-1}(F_1(t)) \leq H_2^{-1}(F_1(t)) + positive.$$

This implies the desired equivalence.

For p.m.'s μ_1 and μ_2 with the same finite mean, it is immediate that $\mu_1 \prec_c \mu_2 \iff \mu_1 \prec_k \mu_2 \iff \mu_1^* \prec_s \mu_2^*$. Under the finite mean assumption, van der Vecht ([22], page 69) gave a result equivalent to the second equivalence and attributed it to Gilat.

Theorem 4.2 *The map* $* : (\mathcal{P}_+, \prec_c) \to (\mathcal{H}^*, \prec_s)$ *is an order-preserving isomorphism.*

Proof: The map $*$ is well defined and one-to-one. Indeed, the one-to-one property follows since $\mu_1^* = \mu_2^*$ implies $F_1^{-1} = F_2^{-1}$ a.e. with respect to the Lebesgue measure, which in turn implies equality of the distribution functions F_1 and F_2. By the previous lemma $*$ is order preserving. q.e.d.

It is an immediate consequence of Theorems 3.4 and 4.2 that $(\mathcal{H}^*, \prec_s)$ is a complete lattice. For the lattice on $\mathcal{H}^*$, the supremum $\vee_*$ and infimum $\wedge_*$ in $\mathcal{H}^*$ are given by

$$\mu_1^* \vee_* \mu_2^* = (\mu_1 \vee_c \mu_2)^* \quad \text{and} \quad \mu_1^* \wedge_* \mu_2^* = (\mu_1 \wedge_c \mu_2)^*.$$

The next example shows that $\vee_s$ can be different from $\vee_*$ on $\mathcal{H}^*$.

Example 4.3 For any point measure, the Hardy-Littlewood maximal p.m. is the same point measure. For two-point measures, if $a < b$ and $0 < \mu(\{b\}) = p = 1 - \mu(\{a\}) = 1 - q < 1$, then

$$F^{-1}(u) = \begin{cases} a & 0 \leq u \leq q \\ b & q < u \leq 1 \end{cases}$$

$$H^{-1}(u) = \begin{cases} (aq + bp - au)/(1-u) & \text{if } 0 < u \leq q \\ b & \text{if } q < u < 1 \end{cases}$$

Now consider a one-point p.m. μ_1 concentrated at c and a two-point p.m. μ_2 concentrated at a and b as above, with $aq + bp < c < b$. Then $H_1^{-1} \vee_s H_2^{-1}$ has associated p.m. $\mu_1^* \vee_s \mu_2^*$ with two points c and b of discontinuity. Therefore $\mu_1^* \vee_s \mu_2^*$ cannot be a Hardy-Littlewood maximal function.

Next we give another isomorphic characterization of $\mathcal{H}^*$ and show that the stochastic infimum in $\mathcal{H}^*$ is the $*$-infimum. These results are based on the following Lemma.

Lemma 4.4 *A function* $G : (0,1) \to I\!R$ *is a Hardy-Littlewood maximal function if and only if* $(1-u)G(u)$ *is a concave function in* $u \in (0,1)$ *and* $\lim_{u \to 1}(1-u)G(u) = 0$.

Proof: We first prove implication '$\Rightarrow$'. Observe that $(1-u)G(u)$ has a derivative which exists a.e. and which is a.e. increasing. This implies that the function $(1-u)G(u)$ is concave. For the other conclusion, observe that $\lim_{u\to 1}(1-u)G(u) = \lim_{u\to 1}\int_u^1 F^{-1}(v)dv = 0$.

Next, we prove the other implication. Let $F^{-1}(u)$ be a left continuous modification of the a.e. derivative of $(1-u)G(u)$. This derivative is increasing and it is easy to show that $(1-u)G(u) = \int_u^1 F^{-1}(u)du$. q.e.d.

Motivated by this Lemma, we introduce the collection J of functions $f : (0,1) \to I\!R$ which are concave and satisfy $\lim_{u\to 1} f(u) = 0$. The usual pointwise order on functions is taken to be the partial order on J, denoted as in Section 3 by $\prec$. To describe the lattice structure on J, let *concave*(S) denote the upper concave envelope of the pointwise supremum of the functions in a collection of real-valued functions S. Define the supremum $\sqcup_0$ and the infimum $\sqcap_0$ on J by

$$(f_1 \sqcup_0 f_2)(t) = (concave(\{f_1, f_2\}))(t) \text{ and } (f_1 \sqcap_0 f_2)(t) = f_1(t) \wedge f_2(t).$$

for all $t \in (0,1)$.

Theorem 4.5 *The spaces* $(\mathcal{H}^*, \prec_s)$ *and* $(J, \prec)$ *are lattice isomorphic.*

Proof: The proof is straightforward using Lemma 4.4. q.e.d.

Corollary 4.6 *The stochastic infimum and the $*$-infimum of a set of Hardy-Littlewood maximal p.m.'s are the same, provided one exists.*

Proof: Let K be a set in $\mathcal{H}^*$, and let ψ map μ^* to the function $(1-u)H_\mu^{-1}(u)$ in $u \in (0,1)$. Then

$$\bigwedge_* K = \bigwedge_* \psi^{-1}(\psi(K)) = \psi^{-1}(\sqcap_0 \psi(K)) = \psi^{-1}(\wedge\psi(K)) = \bigwedge_s K.$$

q.e.d.

Although $\vee_*$ is in general different from $\vee_s$ on $\mathcal{H}^*$, there is one useful instance in which these operations coincide.

Proposition 4.7 *For any chain K in $\mathcal{H}^*$ bounded from above,* $\bigvee_s K = \bigvee_* K$.

Proof: In the last corollary we used that the infimum of concave functions is concave. Here we use that the supremum of increasing concave functions is again concave. The bound on K implies the limit condition.

Let K be the chain in $\mathcal{H}^*$ and ψ map μ^* to the function $(1-u)H_\mu^{-1}(u)$ in $u \in (0,1)$. Then

$$\bigvee_* K = \bigvee_* \psi^{-1}(\psi(K)) = \psi^{-1}(\sqcup_0 \psi(K)) = \psi^{-1}(\vee\psi(K)) = \bigvee_s K.$$

q.e.d.

We give two consequences of these results. The first brings together results which connect the stochastic and the c-convex orders.

Proposition 4.8 *(i) For $\mu_1, \mu_2 \in \mathcal{P}_+$, it follows that*

$$(\mu_1 \wedge_c \mu_2)^* = \mu_1^* \wedge_* \mu_2^* = \mu_1^* \wedge_s \mu_2^* \prec_s \mu_1^* \vee_s \mu_2^* \prec_s \mu_1^* \vee_* \mu_2^* = (\mu_1 \vee_c \mu_2)^*.$$

(ii) For any bounded set K in $\mathcal{H}^*$, it follows that

$$\bigwedge_* K = \bigwedge_s K \prec_s \bigvee_s K \prec_s \bigvee_* K.$$

Proof: All statements follow by previous results, together with the fact that $\bigvee_s K \prec_s \bigvee_* K$. This fact follows by the defining properties for $\bigvee_s$ and $\bigvee_*$. q.e.d.

For use in the next Theorem, we record analogous results for p.m.'s in $\mathcal{P}_-$. For any p.m. μ in $\mathcal{P}_-$ with corresponding d.f. F, define the function $L^{-1} = L_\mu^{-1} : (0,1) \to I\!R$ by $L^{-1}(w) := w^{-1} \int_0^w F^{-1}(u)du$. Associated with this function L^{-1} are its right continuous inverse, a distribution function denoted by L, and the p.m. μ^0. We use $\mathcal{L}^0$ to denote the collection of such p.m.'s μ^0, for $\mu \in \mathcal{P}_-$. From Lemma 4.1 and Theorems 4.2 and 4.4 (or use Lemma 2.5 and Theorem 3.6(a)), it is clear that

(4.1)

(i) For p.m.'s μ_1 and μ_2 in $\mathcal{P}_-$, $\mu_1 \prec_d \mu_2$ is equivalent to $\mu_2^0 \prec_s \mu_1^0$.

(ii) The map $\Box : (\mathcal{P}_-, \prec_d) \to (\mathcal{L}^0, \prec_s)$, defined by $\Box(\mu) = \mu^0$, is an order-reversing isomorphism.

(iii) The space $(\mathcal{L}^0, \prec_s)$ is a complete lattice, with supremum $\vee_0$ and infimum $\wedge_0$ in $\mathcal{L}^0$ given by

$$\mu_1^0 \vee_0 \mu_2^0 = (\mu_1 \wedge_d \mu_2)^0 \text{ and } \mu_1^0 \wedge_0 \mu_2^0 = (\mu_1 \vee_d \mu_2)^0.$$

(iv) A function $K : (0,1) \to I\!R$ has representation $K = L_\mu^{-1}$ for some μ in $\mathcal{P}_-$ if and only if $uK(u)$ is a convex function in $u \in (0,1)$ and $\lim_{u\to 0} uK(u) = 0$.

(v) Let I denote the space of functions $f : (0,1) \to \mathbb{R}$ which are convex and satisfy $\lim_{u\to 0} f(u) = 0$, with the pointwise order on functions $\prec$ and lattice operations $\sqcap$ and $\sqcup$, as described in Section 3. The spaces $(\mathcal{L}^0, \prec_s)$ and $(I, \prec)$ are lattice isomorphic.

(vi) The stochastic supremum and the $\vee_0$-supremum of a set in $\mathcal{L}^0$ coincide, provided one exists; and the stochastic infimum and the $\wedge_0$-infimum of a set in $\mathcal{L}^0$ which is bounded from below coincide.

Theorem 4.9 *Let μ_α, μ_β and μ_1 be p.m.'s with μ_α in $\mathcal{P}_-$, μ_β in $\mathcal{P}_+$, and μ_1 in $\mathcal{P}_r$ such that $\mu_1 \prec_d \mu_\alpha$ and $\mu_1 \prec_c \mu_\beta$. Then the set $\{\mu \in \mathcal{P}(\mathbb{R}) : \mu_1 \prec_k \mu, \mu \prec_d \mu_\alpha$ and $\mu \prec_c \mu_\beta\}$ contains its supremum in the k-convex order. This supremum coincides with the supremum of this set in the c-convex order and in the d-convex order.*

Proof: Let $M := \{\mu \in \mathcal{P}(\mathbb{R}) : \mu_1 \prec_k \mu, \mu \prec_d \mu_\alpha$ and $\mu \prec_c \mu_\beta\}$. The set M is nonempty and is bounded from above in the c-convex order. Since $(\mathcal{P}_+, \prec_c)$ is complete, M has a supremum in the c-convex order, which we denote by γ. It is immediate that $\mu_1 \prec_c \gamma \prec_c \mu_\beta$. We will show that $\mu_1 \prec_k \gamma$ and that γ is also the supremum of M in the k-convex order.

Let D be any dense subset of $\mathbb{R}$. Use diagonalization to obtain a sequence $\{\lambda_n : n = 1,2,\ldots\}$ in M satisfying $\sup_{n\geq 1} \lambda_n^*[x,\infty) = \sup\{\mu^*[x,\infty) : \mu \in M\}$ for every $x \in D$. From the lattice structure on $\mathcal{P}_+$, we may assume that $\{\lambda_n : n = 1,2,\ldots\}$ is increasing in the c-convex order or equivalently $\{\lambda_n^* : n = 1,2,\ldots\}$ is increasing in the stochastic order. Denote $\lambda := \vee_c\{\lambda_n : n = 1,2,\ldots\}$. It is immediate that $\lambda \prec_c \gamma$; we show that $\gamma \prec_c \lambda$. From Proposition 4.7, we have that

$$\lambda^* = (\vee_c\{\lambda_n : n = 1,2,\ldots\})^* = \vee_*\{\lambda_n^* : n = 1,2,\ldots\} = \vee_s\{\lambda_n^* : n = 1,2,\ldots\}.$$

Thus, for each $x \in D$, $\sup\{\mu^*[x,\infty) : \mu \in M\} = \lambda^*[x,\infty)$; and hence λ is an upper bound for M in the c-convex order, and $\gamma \prec_c \lambda$. Thus $\gamma = \lambda = \vee_c\{\lambda_n : n = 1,2,\ldots\}$.

¿From the observation that $\mu_\alpha^0 \prec_s \mu \prec_s \mu_\beta^*$ for each $\mu \in M$, it follows that M is tight. Thus $\{\lambda_n\}$ has a subsequence $\{\lambda_i\}$ which converges weakly to some p.m. ν. Also, M is uniformly integrable. Indeed, the set of p.m.'s M is uniformly integrable from below since

$$\begin{aligned} 0 \leq \lim_{A\to\infty} \sup_{\mu\in M} \int_{x\leq -A} -x d\mu &\leq \lim_{A\to\infty} \sup_{\mu\in M} \int (-2x - A)^+ d\mu \\ &\leq \lim_{A\to\infty} \int (-2x - A)^+ d\mu_\alpha = 0, \end{aligned}$$

where we have used the pointwise inequality $-xI_{(-\infty,-A]}(x) \leq (-2x - A)^+$, Lemma 2.5 $\mu \prec_d \mu_\alpha$ for all $\mu \in M$, and $\mu_\alpha \in \mathcal{P}_-$; and M is uniformly integrable from above from an analogous argument, since $\mu \prec_c \mu_\beta$ for all $\mu \in M$, and $\mu_\beta \in \mathcal{P}_+$.

We show that $\gamma = \nu$. But first we prove that $\nu \in \mathcal{P}_+$. This result uses the fact that $\lim_i \int_{t_0}^{t_1} F_{\lambda_i}^{-1}(u)du = \int_{t_0}^{t_1} F_\nu^{-1}(u)du$ for $0 < t_0 < t_1 < 1$; which follows from (i) $\{F_{\lambda_i}^{-1}(u)\}$ is uniformly bounded over $i = 1, 2, \ldots$ and $u \in [t_0, t_1]$, (ii) $\lim_i F_{\lambda_i}^{-1}(u) = F_\nu^{-1}(u)$ at every continuity point u of F_ν^{-1} in (0,1), and (iii) the bounded convergence theorem. Now, if $\nu \notin \mathcal{P}_+$, then it must be the case that $\lim_{t_1 \nearrow 1} \int_{t_1}^1 F_\nu^{-1}(u)du = \infty$; and thus for any $\epsilon > 0$, for each t_0 close to 1, there is t_1 in $(t_0, 1)$ for which

$$\begin{aligned} \epsilon < \int_{t_0}^{t_1} F_\nu^{-1}(u)du &= \lim_i \int_{t_0}^{t_1} F_{\lambda_i}^{-1}(u)du \\ &\leq \lim_i \int_{t_0}^{1} F_{\lambda_i}^{-1}(u)du = \int_{t_0}^{1} F_\gamma^{-1}(u)du, \end{aligned}$$

where the last equality follows from $\gamma^* = \vee_s\{\lambda_i^* : i = 1, 2, \ldots\}$. This contradicts $\lim_{t_0 \nearrow 1} \int_{t_0}^1 F_\gamma^{-1}(u)du = 0$, i.e., $\gamma \in \mathcal{P}_+$. Thus, it follows that $\nu \in \mathcal{P}_+$. From this result and the uniform integrability M, it follows (e.g., by use of the version of Fatou's Lemma given by Chow and Teicher ([3]: page 94)) that

$$\begin{aligned} \int_w^1 F_\nu^{-1}(u)du &= \int_w^1 \lim_i F_{\lambda_i}^{-1}(u)du \\ &= \lim_i \int_w^1 F_{\lambda_i}^{-1}(u)du = \int_w^1 F_\gamma^{-1}(u)du \end{aligned}$$

for every $w \in (0, 1)$. Thus $\nu^* = \gamma^*$, and so $\nu = \gamma$; the weak convergence limit of $\{\lambda_i\}$ is γ. We now have that $\int x d\gamma = \int x d\lambda_i = \int x d\mu_1$, since

$$\begin{aligned} \int x d\gamma &\geq \lim_i \int x d\lambda_i = \lim_i \int_0^1 F_{\lambda_i}^{-1}(u)du \\ &\geq \int_0^1 \lim_i F_{\lambda_i}^{-1}(u)du = \int_0^1 F_\gamma^{-1}(u)du = \int x d\gamma. \end{aligned}$$

Thus, $\mu \prec_k \gamma$ for all $\mu \in M$; this says that γ is an upper bound for M in the k-convex order. By the completeness of $(\mathcal{P}_r, \prec_k)$ it follows that M has a supremum in the k-convex order; and a straightforward check of the defining properties of suprema shows that γ is also the supremum of M in the k-convex order.

Now, since M is bounded from above in the d-convex order, and $(\mathcal{P}_-, \prec_d)$ is complete, it follows that M has a supremum in the d-convex order, which we call τ. It is immediate that $\mu_1 \prec_d \tau \prec_d \mu_\alpha$. Using the results developed in Lemma 2.5, Theorem 3.6, and (4.1), one shows analogous to the above argument that $\tau = \gamma$, the supremum of M in the k-convex order. It follows also that $\gamma \in M$, and thus the Theorem is proved. q.e.d.

Corollary 4.10 *Let μ_1 and μ_2 be p.m.'s in $\mathcal{P}_+$ with supports in $[0,\infty)$, and assume that $\mu_1 \prec_c \mu_2$. Then the set $M_{kc} = \{\mu \in \mathcal{P}(I\!R) : \mu_1 \prec_k \mu \prec_c \mu_2\}$ contains its supremum in the k-convex order, and this supremum equals the supremum of this set in both the c-convex and d-convex orders.*

Proof: This result follows immediately from Theorem 4.9, observing that (i) p.m. μ has support in $[0,\infty)$ if and only if $\mu \prec_d \epsilon_0$, and (ii) if p.m. μ satisfies $\mu_1 \prec_c \mu \prec_c \mu_2$, then μ has support in $[0,\infty)$. q.e.d.

Example 4.11 In the spirit of Theorem 4.9, one might consider whether the following analogue of a conclusion from Theorem 3.6 holds: if $\mu_1 \in \mathcal{P}(I\!R)$ and $\mu_2 \in \mathcal{P}_r$ with $\mu_1 \prec_c \mu_2$, then $\{\mu \in \mathcal{P}(I\!R) : \mu_1 \prec_k \mu \prec_c \mu_2\}$ has an upper bound in $\mathcal{P}_r$ and contains its supremum in the k-convex order. But this statement is false, and additional assumptions are required as in Theorem 4.9, as the following example illustrates. For $n = 1, 2, \ldots$, denote p.m.'s $\nu_n = n^{-1}\epsilon_{-n} + (1-n^{-1})\epsilon_0$. From the equivalences of Lemmas 2.2 and 2.3, one obtains that $\{\nu_n : n = 1, 2, \ldots\} \subseteq \{\mu \in \mathcal{P}(I\!R) : \epsilon_{-1} \prec_k \mu \prec_c \epsilon_0\} =: B$. However, the set $\{\nu_n : n = 1, 2, \ldots\}$ has no upper bound within $\mathcal{P}_{-1}$ in the k-convex order; and thus the set B has no upper bound within $\mathcal{P}_{-1}$ in the k-convex order. We also have that B has no upper bound in $\mathcal{P}_-$ within the d-convex order. (Similarly, by considering the p.m.'s $\{\xi_n = n^{-1}\epsilon_n + (1-n^{-1})\epsilon_0 : n = 1, 2, \ldots\}$, one sees that the set of p.m.'s $\{\mu \in \mathcal{P}(I\!R) : \epsilon_1 \prec_k \mu \prec_d \epsilon_0\}$ might have no upper bound within $\mathcal{P}_1$ in the k-convex order and no upper bound within $\mathcal{P}_+$ in the c-convex order.)

5 Connections to Martingale Theory

In this Section we give some connections of the orders $\prec_c$, $\prec_k$, and $\prec_d$ to martingale theory.

We say that r.v. X is smaller than r.v. Y in the stochastic order, in the c-convex order or in the k-convex order, if the corresponding p.m.'s μ_1 and μ_2 satisfy respectively $\mu_1 \prec_s \mu_2$, $\mu_1 \prec_c \mu_2$ or $\mu_1 \prec_k \mu_2$. We shall use the same notation for distribution functions with the obvious meaning. A *martingale* (resp., *submartingale or supermartingale*) is a collection of r.v.'s X_t, $t \in T \subset I\!R$ and an increasing family of σ-fields $\mathcal{F}_t \subset \mathcal{F}$, $t \in T$, on the same probability space $(\Omega, \mathcal{F}, P)$ satisfying the following: (i) the r.v. X_t is $\mathcal{F}_t$ measurable, for each $t \in T$, and the map $X_.(\omega) : T \to I\!R$ is right continuous *a.e.* P; (ii) $E(X_t)$ is well-defined (but allowed to be finite,$-\infty$, or $+\infty$; and (iii) $E(X_t|\mathcal{F}_s) =$ (resp. $\geq$, or $\leq$)X_s *a.e.* P for $s \leq t$ in T. We call (μ_1, μ_2) a *martingale pair* (resp., *submartingale pair or supermartingale pair*), if there is a martingale (resp., submartingale or supermartingale) X_1, X_2 for which X_1 and X_2 have distributions μ_1 and μ_2 respectively. Martingale, submartingale, and supermartingale triples (μ_1, μ_2, μ_3), quadruples

$(\mu_1, \mu_2, \mu_3, \mu_4)$, etc. are defined analogously. A *probability kernel* is a map $k : \Omega \text{x} \mathcal{F} \rightarrow [0,1]$ with (i) $k(\omega, .) : \mathcal{F} \rightarrow [0,1]$ is a probability measure *a.e.* P, and (ii) $k(., F) : \Omega \rightarrow [0,1]$ is measurable for all $F \in \mathcal{F}$. A *dilation* is a probability kernel k for which $x = \int y k(x, dy)$ for all $x \in \mathbb{R}$. The next theorem is partly due to Strassen and others, as indicated in [21].

Theorem 5.1 *Let μ_1 and μ_2 be p.m.'s in $\mathcal{P}_r$. Then the following are equivalent:*

(i) $\mu_1 \prec_k \mu_2$.

(ii) (μ_1, μ_2) *is a martingale pair.*

(iii) There exists a dilation $k : \mathbb{R} \text{x} \mathcal{B} \rightarrow [0,1]$ with $\int k(x, B) d\mu_1(x) = \mu_2(B)$ for all Borel sets $B \in \mathcal{B}$.

We give here a slightly stronger result adapted to $\mathcal{P}_+$, which generalizes a finite moment submartingale characterization given in [21].

Theorem 5.2 *Let μ_1 and μ_2 be p.m.'s in $\mathcal{P}_+$. Then the following are equivalent:*

(i) $\mu_1 \prec_c \mu_2$.

(ii) (μ_1, μ_2) *is a submartingale pair.*

(iii) There exists a probability kernel $k : \mathbb{R} \text{x} \mathcal{B} \rightarrow [0,1]$ with $x \leq \int y k(x, dy)$ for all $x \in \mathbb{R}$ and $\int k(x, B) d\mu_1(x) = \mu_2(B)$ for all Borel sets $B \in \mathcal{B}$.

Proof: The equivalence of (ii) and (iii) is straightforward. In particular, for (ii) $\Rightarrow$ (iii), if X_1 and X_2 are r.v.'s with associated p.m.'s μ_1 and μ_2, and $\{X_1, X_2\}$ is a submartingale, then a regular conditional distribution of X_2 given X_1 can be taken as the kernel in (iii). For (iii) $\Rightarrow$ (ii), on $(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2))$, define p.m. λ by $\lambda(A \times B) = \int_A k(x, B) d\mu_1(x)$, for A, B in $\mathcal{B}(\mathbb{R})$; define r.v.'s X_1 and X_2 as the coordinate r.v.'s $(X_1(x,y), X_2(x,y)) = (x,y)$; and define σ-fields $\mathcal{F}_1 = \mathcal{F}(X_1)$, the σ-field generated by X_1, and $\mathcal{F}_2 = \mathcal{F}(X_1, X_2)$. It is immediate that $\{X_1, X_2\}$ is a submartingale with marginals μ_1 and μ_2, and hence that (μ_1, μ_2) is a submartingale pair.

The implication that (ii) $\Rightarrow$ (i) follows from the generalized Jensen's inequality (e.g., see Chow and Teicher [3]: Theorem 7.1.4). We prove that (i) $\Rightarrow$ (ii) and (iii). The additional difficulty of the proof of the implication that (i) $\Rightarrow$ (ii) over that of Theorem 5.1 is the one-sided integrability of the measures. We obtain this result by a reduction to the setting of Theorem 5.1. For simplicity in exposition, we give the proof in the case of a strictly increasing and continuous d.f. F. For the general d.f. F, one uses the basic outline of this proof together with the 'splitting atom' technique, as used e.g., in Proposition 2.3 of [12].

We saw in Lemma 2.2 that $\mu_1 \prec_c \mu_2$ is equivalent to $\int_w^1 F_1^{-1}(u)du \le \int_w^1 F_2^{-1}(u)du$ for all $w \in (0,1)$. The function $g : (0,1) \to I\!R$ defined by $g(u) := (F_2^{-1} - F_1^{-1})(u)$ is left continuous. If g is always positive then $\mu_1 \prec_s \mu_2$ and we are done. Otherwise let u_0 be a value with $g(u_0) < 0$. Define $u_1 := \sup\{u < u_0 : \ g(u) \ge 0\}$ and $u_2 := \inf\{u > u_0 : g(u) \ge 0\}$. Note that u_1 might be zero and that $\int_{u_1}^{u_2} g(u)du$ is finite. Then for all $u \in (u_1, u_2)$ it follows that $g(u) < 0$. Further choose some u_4 with $u_1 < u_2 < u_4$ and

$$\int_{u_2}^{u_4} g^+(u)du = -\int_{u_1}^{u_2} g(u)du.$$

This is possible by $\int_w^1 g(u)du \ge 0$ for all $w \in (0,1)$.

Consider $g(u)du$ as a signed measure μ on $(0,1)$. (Note that $|\mu(A)| = |\int_A g(u)du|$ is finite for any Borel set $A \subset (w,1)$, for some $w \in (0,1)$, and $0 \le \mu(0,1) \le \infty$.) Take the Hahn-Jordan decomposition $\mu = \mu^+ - \mu^-$ into positive and negative part μ^+, μ^-. Define $I := ([u_2,u_4] \cap \mathit{support}(\mu^+)) \cup ([u_1,u_2] \cap \mathit{support}(\mu^-)) = ([u_2,u_4] \cap \mathit{support}(\mu^+)) \cup [u_1,u_2]$. For any p.m. ν on the Borel sets of (0,1), let ν_I denote the restriction to set I, i.e., $\nu_I(\cdot) = \nu(\cdot)/\nu(I)$ on the Borel subsets of I. Let G_1, G_2 be the distribution functions for $\mu_{1F_1^{-1}(I)}, \mu_{2F_2^{-1}(I)}$ respectively. Then we claim

1) $\mu_{1F_1^{-1}(I)}(I\!R) = \mu_{2F_2^{-1}(I)}(I\!R)$

2) $F_1(G_1^{-1}(w)) = F_2(G_2^{-1}(w))$ for all $w \in (0,1)$.

3) $\mu_{1F_1^{-1}(I)} \prec_c \mu_{2F_2^{-1}(I)}$

4) $\mu_{1F_1^{-1}(I)} \prec_k \mu_{2F_2^{-1}(I)}$

5) $\mu_{1F_1^{-1}(I^c)} \prec_c \mu_{2F_2^{-1}(I^c)}$

For 1), observe $\mu_{1F_1^{-1}(I)}(I\!R) = \int_I du = \mu_1(F_1^{-1}(I) \cap I\!R)$.

For 2), one shows that $G_1(F_1^{-1}(v)) = \mu_1((-\infty, F_1^{-1}(v)) \cap F_1^{-1}(I)) / \int_I du = \int_{(0,v)\cap I} du / \int_I du$. Also $G_2(F_2^{-1}(v)) = \int_{(0,v)\cap I} du / \int_I du$. This implies the statement 2).

For 3), observe that

$$\begin{aligned} &\mu_{1F_1^{-1}(I)} \prec_c \mu_{2F_2^{-1}(I)} \\ \iff &\int_w^1 G_1^{-1}(u)du \le \int_w^1 G_2^{-1}(u)du \text{ for all } w \\ \iff &\int_{G_1^{-1}(w)}^{\infty} x dG_1(x) \le \int_{G_2^{-1}(w)}^{\infty} x dG_2(x) \text{ for all } w \end{aligned}$$

$$\iff \int_{(G_1^{-1}(w),\infty)\cap F_1^{-1}(I)} x dF_1(x) \le \int_{(G_2^{-1}(w),\infty)\cap F_2^{-1}(I)} x dF_2(x) \text{ for all } w$$

$$\iff \int_{(F_1(G_1^{-1}(w)),1)\cap I} F_1^{-1}(u)du \le \int_{(F_2(G_2^{-1}(w)),1)\cap I} F_2^{-1}(u)du \text{ for all } w$$

$$\iff \int_{(w,1)\cap I} F_1^{-1}(u)du \le \int_{(w,1)\cap I} F_2^{-1}(u)du \text{ for all } w.$$

Now consider $\int_{(w,1)\cap I} g(u)du$. Notice $\int_I g(u)du = 0$. The integral $\int_{(w,1)\cap I} g(u)du$ equals 0 for $w \ge u_4$; increases from 0 to $\int_{[u_2,u_4]\cap \mathrm{supp}(\mu^+)} g(u)du = \int_{[u_2,u_4]} g^+(u)du$ as w decreases from u_4 to u_2; decreases from $\int_{[u_2,u_4]} g^+(u)du$ to 0 as w decreases from u_2 to u_1; and equals 0 for $w \le u_1$. This proves the statement 3).

Statement 4) follows since by construction both measures μ_1, μ_2 restricted to $F_1^{-1}(I), F_2^{-1}(I)$ respectively have the same finite first moment equal to $\int_I F_1^{-1}(u)du/\int_I du = \int_I F_2^{-1}(u)du/\int_I du$.

By reasoning analogous to that used in claims 1)–3), one reduces verification of claim 5) to showing $\int_{(w,1)\cap I^c} g(u)du \ge 0$ for all w. For $w \ge u_4$ this expression is positive. For $u_1 \le w \le u_4$ establish first $\int_{(u_1,1)\cap I^c} g(u)du \le \int_{(w,1)\cap I^c} g(u)du$. Then argue

$$0 \le \int_{(u_1,1)} g(u)du = \int_{(u_1,1)\cap I^c} g(u)du + \int_{(u_1,1)\cap I} g(u)du = \int_{(u_1,1)\cap I^c} g(u)du.$$

For $w \le u_1$ argue $0 \le \int_w^1 g(u)du = \int_{(w,1)\cap I^c} g(u)du$.

By Theorem 5.1 there is a martingale pair for μ_1, μ_2 restricted to $F_1^{-1}(I)$, $F_2^{-1}(I)$ respectively.

Now continue the procedure for $\mu_1 \prec_c \mu_2$ restricted to $F_1^{-1}(I^c), F_2^{-1}(I^c)$. We find a new set I_2, disjoint from $I = I_1$ other than for possibly countably many points, by the above procedure. Again there is a martingale pair for μ_1, μ_2 restricted to $F_1^{-1}(I_2), F_2^{-1}(I_2)$ respectively. We are left with $\mu_1 \prec_c \mu_2$ restricted to the inverse of $(I_1 \cup I_2)^c$ under F_1, F_2. We finally have to show that this algorithm will work to show that (μ_1, μ_2) is a submartingale pair. Let D be a countable dense set in $\{u \in (0,1) : g(u) < 0\}$. Take an enumeration of D and work through this construction taking the n^{th} element as starting point in the n^{th} step if necessary. In at most countably many steps, one obtains a sequence of measurable sets $\{I_n\}_{1\le n<\infty}$ disjoint except possibly for at most countably many points, with $\{u \in (0,1) : g(u) < 0\} \subset \bigcup_{1\le n<\infty} I_n$, and $(\mu_{1F_1^{-1}(I_n)}, \mu_{2F_2^{-1}(I_n)})$ is a martingale pair for each $1 \le n < \infty$. On $(\cup_n I_n)^c$, we have that $\mu_{1F_1^{-1}((\cup_n I_n)^c)} \prec_s \mu_{2F_2^{-1}((\cup_n I_n)^c)}$; and so μ_1, μ_2 restricted to $F_1^{-1}((\cup_n I_n)^c), F_2^{-1}((\cup_n I_n)^c)$ is a submartingale pair. For $1 \le j < \infty$, let k_j be a dilation for $\mu_{1F_1^{-1}(I_j)}, \mu_{2F_2^{-1}(I_j)}$; and let k_∞ be a probability kernel associated with submartingale pair μ_1, μ_2 restricted to $F_1^{-1}((\cup_n I_n)^c), F_2^{-1}((\cup_n I_n)^c)$. For $i = 1, 2$, denote sets $\Omega_{i1} = F_i^{-1}(I_1)$; $\Omega_{ij} = F_i^{-1}(I_j) \cap F_i^{-1}((\cup_{\ell=1}^{j-1} I_\ell)^c)$ for $1 < j < \infty$; and $\Omega_{i\infty} = F_i^{-1}((\cup_n I_n)^c)$. Define k on $\mathbb{R} \times \mathcal{B}$ by $k(x, B) = \sum_{1 \le j \le \infty} k_j(x, B) I_{\Omega_{1j}}(x)$. Then $k(\cdot, B)$ if a Borel measurable function, for each Borel set B in $\mathbb{R}$; and for each $x \in (0, 1)$, there is some $j \in \{1, \ldots, \infty\}$ for each $k(x, \cdot) = k_j(x, \cdot)$, and so $k(x, \cdot)$ is a p.m. on the Borel sets of $\mathbb{R}$. So k is a probability kernel. Finally, observe that

$$\begin{aligned} \int yk(x,dy) &= \int yk_j(x,dy) = x \qquad \text{if } x \in \Omega_{1j}, \text{ for } 1 \le j < \infty; \text{ and} \\ &= \int yk_\infty(x,dy) \ge x \qquad \text{if } x \in \Omega_{1\infty}; \end{aligned}$$

and for each Borel set B in $\mathbb{R}$,

$$\begin{aligned} \int_{\mathbb{R}} k(x,B)d\mu_1(x) &= \sum_{1 \le j \le \infty} \int_{\Omega_{1j}} k_j(x,B)d\mu_1 = \sum_{1 \le j \le \infty} \int k_j(x,B)d\mu_{1\Omega_{1j}} \\ &= \sum_{1 \le j \le \infty} \mu_{2\Omega_{2j}}(B) = \sum_{1 \le j \le \infty} \mu_2(B \cap \Omega_{2j}) = \mu_2(B) \end{aligned}$$

Thus, the kernel k satisfies (iii) for p.m.'s μ_1 and μ_2; and the implication (i) $\Rightarrow$ (ii) and (iii) is proved. (An alternative proof can be given based on Zorn's Lemma; however, the proof we have given here has the advantage of a countable construction.) q.e.d.

In the martingale theory terminology of this Section, part of Theorem 3.5 can be restated as the following result.

Theorem 5.3 *Let μ and ν be p.m.'s in $\mathcal{P}_+$ and $\mathcal{P}_r$ respectively for which (μ, ν) is a submartingale pair. Then there is a p.m. $\mu_1 \in \mathcal{P}_r$ for which (i) (μ_1, ν) is a martingale pair; (ii) (μ, μ_1, ν) is a submartingale triple; and (iii) there is no p.m. μ_0 in $\mathcal{P}_r$ with $\mu_0 \ne \mu_1$ satisfying (μ_0, μ_1, ν) is a martingale triple and (μ, μ_0, μ_1, ν) is a submartingale quadruple.*

Here is an analogue of Theorem 5.2 for supermartingales which we use in Theorem 5.5.

Theorem 5.4 *Let μ_1 and μ_2 be p.m.'s in $\mathcal{P}_-$. Then the following are equivalent:*

(i) $\mu_1 \prec_d \mu_2$.

(ii) (μ_1, μ_2) *is a supermartingale pair.*

(iii) There exists a probability kernel $k : \mathbb{R} \times \mathcal{B} \to [0, 1]$ with $x \ge \int yk(x, dy)$ for all $x \in \mathbb{R}$ and $\int k(x, B)d\mu_1(x) = \mu_2(B)$ for all Borel sets $B \in \mathcal{B}$.

In the martingale theory terminology of this Section, Theorem 4.9 can be restated to give this variation of Theorem 5.3.

Theorem 5.5 *Let μ, ν_α and ν_β be p.m.'s in $\mathcal{P}_r, \mathcal{P}_-$ and $\mathcal{P}_+$ respectively for which (μ, ν_α) is a submartingale pair and (μ, ν_β) is a supermartingale pair. Then there is a p.m. $\mu_1 \in \mathcal{P}_r$ for which*

(i) (μ, μ_1) *is a martingale pair;*

(ii) (μ, μ_1, ν_α) *is a submartingale triple and* (μ, μ_1, ν_β) *is a supermartingale triple;*

(iii) there is no *p.m.* μ_2 *in* $\mathcal{P}_r$ *with* $\mu_2 \neq \mu_1$ *satisfying* (μ, μ_1, μ_2) *is a martingale triple and either* $(\mu, \mu_1, \mu_2, \mu_\alpha)$ *is a submartingale quadruple or* $(\mu, \mu_1, \mu_2, \mu_\beta)$ *is a supermartingale quadruple.*

Acknowledgment. The authors wish to thank the editors for the opportunity to contribute to this Volume that honors the eminent Thomas Ferguson, and wish to thank the referees for comments that improved the exposition of this paper.

References

[1] Blackwell, D. and Dubins, L.E. (1963) *A converse to the dominated convergence theorem.* Illinois J. Math. 7, 508-514.

[2] Choquet, G. (1969) *Lectures on Analysis.* Vol. II, Benjamin, New York.

[3] Chow, Y. S. and Teicher, H. (1978) *Probability Theory: Independence, Interchangeability, Martingales.* Springer-Verlag. New York.

[4] Dubins, L.E. and Gilat, D. (1978) *On the distribution of maxima of martingales.* Proc. A.M.S. Soc. 68, 337-338.

[5] Durrett, R. (1984) *Brownian Motion and Martingales in Analysis.* Wadsworth, Belmont, California.

[6] Elton, J. and Hill, T.P. *Fusions of a probability distribution.* Ann. Probability 20, 421-454.

[7] Gilat, D. (1987) *On the best order of observation in optimal stopping problems.* J. Apppl. Prob. 24, 773-778.

[8] Gleason,A.M. (1966) *Fundamentals of Abstract Analysis.* Addison-Wesley, London.

[9] Hardy, G.H. and Littlewood, J.E. (1930) *A maximal theorem with function theoretic applications.* Acta. Math. 54, 81-116.

[10] Hill, T.P. and Kertz, R.P. (1983) *Stop rule inequalities for uniformly bounded sequences of random variables.* Trans. A.M.S. 278, 197-207.

[11] Karlin, S. and Novikov, A. (1963) *Generalized convex inequalities.* Pacific J. Math. 13, 1251-1279.

[12] Kertz, R.P. and Rösler, U. (1990) *Martingales with given maxima and terminal distributions.* Israel Journ. Math. 69, 173-192.

[13] Kertz, R.P. and Rösler, U. (1991) *Stochastic and convex orders and lattices of probability measures, with a martingale application.* Israel Journ. Math. 77, 129-164.

[14] Krengel, U. and Sucheston, L. (1977) *Semiamarts and finite values.* Bulletin Amer. Math. Soc., 83, 745-747.

[15] Lehman, E.L. (1959) *Testing statistical hypothesis.* John Wiley & Sons, New York.

[16] Levy, H. (1990) *Stochastic dominance and expected utility: survey and analysis.* Management Sci. 38, 555-593.

[17] Marshall, A.W. and Proschan, F. (1970) *Mean life of series and parallel systems.* J. Appl. Prob. 7, 165-174.

[18] Meyer, P.A. (1966) *Probability and Potentials.* Blaisdell, London.

[19] Phelps, R.R. (1966) *Lectures on Choquet's Theorem.* ed.Halmos, P.R. and Gehring, F.W., van Nostrand Mathematical Studies No 7. Princeton University Press, Princeton.

[20] Stoyan, D. (1983) *Comparison Methods for Queues and other Stochastic Models.* Edited by D.J. Daley. John Wiley & Sons, New York.

[21] Strassen, V. (1965) *The existence of probability measures with given marginals.* Ann. Math. Statist. 36, 423-439.

[22] van der Vecht, D. P. (1986) *Inequalities for Stopped Brownian Motion.* C.W.I. Tract 21, Mathematisch Centrum, Amsterdam.