ISRAEL JOURNAL OF MATHEMATICS 77 (1992), 129–164

STOCHASTIC AND CONVEX ORDERS AND LATTICES OF PROBABILITY MEASURES, WITH A MARTINGALE INTERPRETATION

BY

ROBERT P. KERTZ*

School of Mathematics, Georgia Institute of Technology
Atlanta, Georgia 30332, USA

AND

UWE RÖSLER**

Institut für Mathematische Statistik und Wirtschaftsmathematik
Universität Göttingen, Lotzestrasse 13, 3400 Göttingen, Germany

ABSTRACT

Let μ be any probability measure on $\mathbb{R}$ with $\int |x| d\mu(x) < \infty$ and let μ^* denote the associated Hardy and Littlewood maximal p.m., the p.m. of the Hardy and Littlewood maximal function obtained from μ. Dubins and Gilat [6] showed that μ^* is the least upper bound, in the usual stochastic order, of the collection of p.m.'s ν on $\mathbb{R}$ for which there is a martingale $(X_t)_{0 \le t \le 1}$ having distributions of X_1 and $\sup_{0 \le t \le 1} X_t$ given by μ and ν respectively. In this paper, a type of 'dual representation' is given. Specifically, let ν be any p.m. on $\mathbb{R}$ with $\limsup_{x \to \infty} x\nu[x, \infty) = 0$ and finite $x_0 = \inf\{z : \nu(-\infty, z]\, 0\}$. Then there is a 'minimal p.m.' ν_Δ which is the greatest lower bound, in the usual convex order, of the collection of p.m.'s μ on $\mathbb{R}$ for which there is a martingale $(X_t)_{0 \le t \le 1}$ having distributions of X_1 and $\sup_{0 \le t \le 1} X_t$ given by μ and ν respectively. To demonstrate existence and to obtain identification of these minimal p.m.'s, we use, in particular, a lattice structure on the set of p.m.'s with the convex order, and an equivalence between a convex order of p.m.'s and the stochastic order of their maximal p.m.'s. Consequences of these order results include sharp

* Supported in part by NSF grants DMS-86-01153 and DMS-88-01818.

** *Current address:* Department of Mathematics, Uppsala University, Thunbergsvägen 3, S-752 38 Uppsala, Sweden.

Received March 17, 1991

expectation-based inequalities for martingales. These martingale inequalities form a new class of 'prophet inequalities' in the context of optimal stopping theory.

Introduction

Stochastic and convex orders on probability measures have been used extensively to express comparisons of random variables which describe some feature common to stochastic processes from a given class, such as waiting times for queueing systems with given interarrival time and service time distributions; system lifetimes for reliability systems of the same type, with given component lifetimes; and rates of return on stocks or mutual funds from a given group (see e.g. [14, 9] and references therein). In probabilistic potential theory, convex orders on measures are present in the theory of balayage or 'sweeping out' of one measure to another, and lead to results such as Choquet's integral representation theorems, Skorokhod embedding theorems, and Strassen's martingalizability theorems (see e.g., [5, 10, 15] and references therein). In this paper, we compare terminal elements of martingales having the same maximum distribution, and find that martingale with minimal terminal element in the convex order; thus, we exhibit a kind of 'constrained balayage' comparison.

This is a type of 'dual investigation' to the use of the stochastic order on probability measures to draw a precise connection between the maxima of martingales having the same terminal element. In both investigations, an essential role is played by the Hardy and Littlewood maximal p.m. μ^*, the probability measure of the Hardy and Littlewood maximal function associated with a p.m. μ for which $\int_0^\infty x d\mu(x) < \infty$ (for a definition, see the remarks preceding Lemma 1.8). The comparison of maxima results are given in [1, 6, 8], and yield, in particular, that for any p.m. μ on $\mathbb{R}$ with $\int |x| d\mu(x) < \infty$,

(0.1)

(i) there is a martingale $(Z_t)_{0\le t\le 1}$ for which $Z_1 \stackrel{\mathcal{D}}{=} \mu$ and $\sup_{0\le t\le 1} Z_t \stackrel{\mathcal{D}}{=} \mu^*$;

(ii) $\mu * = \text{l.u.b.}_{\prec_s} \{\nu :$ there is a martingale $(X_t)_{0\le t\le 1}$ satisfying $X_1 \stackrel{\mathcal{D}}{=} \mu$ and $\sup_{0\le t\le 1} X_t \stackrel{\mathcal{D}}{=} \nu\}$;

(iii) $\{\nu :$ there is a martingale $(X_t)_{0\le t\le 1}$ satisfying $X_1 \stackrel{\mathcal{D}}{=} \mu$ and

$\sup_{0\le t\le 1}\{X_t \overset{\mathcal{D}}{=} \nu\} = \{\nu \text{ is a p.m. on } \mathbb{R} : \mu \prec_s \nu \prec_s \mu^*\}.$
(Here $\prec_s$ denotes the stochastic order and $X \overset{\mathcal{D}}{=} \mu$ denotes that X has associated p.m. μ.)

In the development leading to the results on comparison of terminal elements in this paper, the underlying analysis uses a lattice structure on probability measures in the convex order, and an equivalence between convex ordered p.m.'s and stochastic ordered maximal p.m.'s. These lattice and equivalence results are given in Section 1; they are connected to results in [7, 14, 16]. Just as the Hardy and Littlewood maximal p.m. μ^* is associated to any p.m. μ satisfying a right-tail integrability assumption, we find that there is an appropriate 'minimal p.m.' ν_Δ associated to any p.m. ν satisfying a right-tail growth condition. Specifically, we prove the following.

THEOREM A: *Let ν be any p.m. on $\mathbb{R}$ satisfying* $\limsup_{x\to\infty} x\nu[x,\infty) = 0$, *and denote*

$$\mathcal{E}_+(\nu) = \left\{ \mu \text{ is a p.m. on } \mathbb{R} \text{ satisfying} \int_0^\infty x\,d\mu(x) < \infty \text{ and } \nu \prec_s \mu^* \right\}.$$

Then $\mathcal{E}_+(\nu)$ is nonempty and there is a unique p.m. ν_Δ in $\mathcal{E}_+(\nu)$ which is the greatest lower bound of $\mathcal{E}_+(\nu)$ in the convex order.

The proof of Theorem A is immediate from Proposition 2.1 and Theorem 2.4. The p.m. ν_Δ of Theorem A is called the minimal probability measure associated with p.m. ν. Examples of minimal p.m.'s are developed through Lemma 2.5 and Proposition 2.7. Minimal p.m.'s are used in Theorem 3.1 to link the lattice structure on probability measures in the convex order to a lattice structure on maximal p.m.'s in the stochastic order. Some important properties of minimal p.m.'s are given in Proposition 4.1 and Theorem 4.5.

The results which give a comparison of terminal variables are presented in Section 5, and yield in particular the following counterpart to the results of (0.1), which holds for any p.m. ν on $\mathbb{R}$ satisfying $\limsup_{x\to\infty} x\nu[x,\infty) = 0$ with finite $x_0 := \inf\{z : \nu(-\infty, z] > 0\}$,

(0.2)

(i) there is a martingale $(W_t)_{0\le t\le 1}$ for which
$W_1 \overset{\mathcal{D}}{=} \nu_\Delta$ and $\sup_{0\le t\le 1} W_t \overset{\mathcal{D}}{=} \nu$;

(ii) $\nu_\Delta = \text{g.l.b.}_{\prec_c}\{\nu$: there is a martingale $(X_t)_{0\le t\le 1}$ satisfying

$X_1 \stackrel{\mathcal{D}}{=} \mu$ and $\sup_{0\le t\le 1} X_t \stackrel{\mathcal{D}}{=} \nu\}$; and

(iii) $\{\mu$: there is a martingale $(X_t)_{0\le t\le 1}$ satisfying $X_1 \stackrel{\mathcal{D}}{=} \mu$ and $\sup_{0\le t\le 1} X_t \stackrel{\mathcal{D}}{=} \nu\} = \{\mu$ is a p.m. on $\mathbb{R} : \int x d\mu(x) = x_0$ and $\nu_\Delta \prec_c \mu \prec_s \nu\}$.

(Here $\prec_c$ denotes the convex order.) The results of (0.2) follow immediately from Theorem 5.1 and Corollary 5.2. As a consequence of these comparison results, new sharp, expectation-based inequalities for martingales are given in Theorem 5.3. In Remark 5.4, these martingale inequalities are interpreted as a new class of 'prophet inequalities' in optimal stopping theory. This 'prophet vs. gambler' interpretation lies at the heart of the p.m. comparison results in this paper, and was a constant source of motivation throughout the development of this investigation.

1. Partial Orders and Lattices on Sets of Probability Measures

In this section, we give properties of three partial orders $\prec_s, \prec_c$, and $\prec_k$ defined respectively on spaces

$$(1.1)\quad \begin{aligned} \mathcal{P}(\mathbb{R}) &= \text{ the collection of probability measures on } \mathbb{R} = (-\infty,\infty), \\ \mathcal{E}_+ &= \{\mu \in \mathcal{P}(\mathbb{R}) : \int_0^\infty x d\mu(x) < \infty\}, \text{ and} \\ \mathcal{E} &= \{\mu \in \mathcal{P}(\mathbb{R}) : \int |x| d\mu(x) \text{ is finite }\}. \end{aligned}$$

We also discuss $(\mathcal{P}(\mathbb{R}), \prec_s), (\mathcal{E}_+, \prec_c)$, and $(\mathcal{E}, \prec_k)$ as lattices.

For each p.m. μ on $\mathbb{R}, F = F_\mu$ denotes the distribution function of μ and $F^{-1} = F_\mu^{-1}$ denotes its left inverse (left-continuous inverse). F^{-1} is defined on (0,1) by $F^{-1}(w) = \inf\{z : F(z) \ge w\}$, and is extended to [0,1] by setting $F^{-1}(0) = F^{-1}(0+)$ and $F^{-}(1) = F^{-1}(1-)$. In this way, F^{-1} becomes a (extended) random variable on $([0,1], \mathcal{B}([0,1]), m)$ where $\mathcal{B}([0,1])$ denotes the Borel sets on $[0,1]$ and $m(dt) = dt$ denotes Lebesgue measure.

The order $\prec_s$ is defined on $\mathcal{P}(\mathbb{R})$ by $\mu_1 \prec_s \mu_2$ iff $\mu_1[x,\infty) \le \mu_2[x,\infty)$ for all $x \in \mathbb{R}$. Then $\prec_s$ is a partial order on $\mathcal{P}(\mathbb{R})$ and, from [14], one obtains the following.

LEMMA 1.1: *Each of the following is equivalent to $\mu_1 \prec_s \mu_2$ for μ_1 and μ_2 in $\mathcal{P}(\mathbb{R})$:*

(i) $\int \phi d\mu_1 \leq \int \phi d\mu_2$ *for all nondecreasing functions* ϕ *for which both integrals exist; and*
(ii) $F_{\mu_1}^{-1}(w) \leq F_{\mu_2}^{-1}(w)$ *for all* $w \in [0,1]$.

Observe that $\mathcal{P}(\mathbb{R})$ is a lattice in the $\prec_s$ order, under the following operators $\vee_s$ and $\wedge_s$. If ν_1 and ν_2 are in $\mathcal{P}(\mathbb{R})$ with respective r.v.'s F_1^{-1} and F_2^{-1} on $([0,1], \mathcal{B}([0,1]), m)$, then $\nu_1 \vee_s \nu_2$ is the p.m. associated with r.v. $F_1^{-1} \vee F_2^{-1}$, and $\nu_1 \wedge_s \nu_2$ is the p.m. associated with r.v. $F_1^{-1} \wedge F_2^{-1}$.

Next, we define the order $\prec_c$ on the set of p.m.'s $\mathcal{E}_+$ of (1.1) by $\mu_1 \prec_c \mu_2$ iff $\int_t^\infty \mu_1(x,\infty)dx \leq \int_t^\infty \mu_2(x,\infty)dx$ for all $t \in \mathbb{R}$. Then $\prec_c$ is a partial order on $\mathcal{E}_+$ [14], and from [14] one has the following.

LEMMA 1.2: *Each of the following is equivalent to* $\mu_1 \prec_c \mu_2$ *for p.m.'s* μ_1 *and* μ_2 *in* $\mathcal{E}_+$:
(i) $\int (x-t)_+ d\mu_1(x) \leq \int (x-t)_+ d\mu_2(x)$ *for all* $t \in \mathbb{R}$;
(ii) $\int \phi d\mu_1 \leq \int \phi d\mu_2$ *for all nondecreasing convex functions* ϕ *for which the integrals exist.*

We show that $(\mathcal{E}_+, \prec_c)$ is a lattice by finding an appropriate isomorphism between $\mathcal{E}_+$ and the space $\mathcal{B}_+$ of functions on $\mathbb{R}$, defined as the collection of functions ψ on $\mathbb{R}$ satisfying

$$(1.2) \qquad \begin{aligned} &\text{(i) } \psi \text{ is convex and nondecreasing,} \\ &\text{(ii) } \lim_{t\uparrow\infty}(\psi(t)-t)=0, \text{ and} \\ &\text{(iii) } \lim_{t\downarrow-\infty}(\psi(t+h)-\psi(t))=0 \text{ for all } h\in\mathbb{R}. \end{aligned}$$

The usual pointwise order on functions is taken to be the partial order on $\mathcal{B}_+$, denoted $\prec$. We first identify the lattice structure on $(\mathcal{B}_+, \prec)$ and then carry this structure over to $(\mathcal{E}_+, \prec_c)$. To do this, first recall from ([13]: page 37) that the **convex hull of functions** $\{f_i : i \in I\}$ on $\mathbb{R}$, denoted $\mathrm{conv}(\{f_i : i \in I\})$, is the convex hull of the pointwise infimum of $\{f_i : i \in I\}$. It is the greatest convex function f on $\mathbb{R}$ such that $f(x) \leq f_i(x)$ for every $x \in \mathbb{R}$ and every $i \in I$.

LEMMA 1.3: *The space* $(\mathcal{B}_+, \prec)$ *is a lattice under operations* $\vee$ *and* $\wedge$, *defined for* ψ_1 *and* ψ_2 *in* $\mathcal{B}_+$ *by*

$$(1.3) \quad (\psi_1 \vee \psi_2)(t) = \psi_1(t) \vee \psi_2(t) \quad \text{and} \quad (\psi_1 \wedge \psi_2)(t) = (\mathrm{conv}(\{\psi_1, \psi_2\}))(t).$$

Proof: Let ψ_1 and ψ_2 be in $\mathcal{B}_+$ and define $\psi_1 \vee \psi_2$ and $\psi_1 \wedge \psi_2$ as in (1.3). It is straightforward to show that $\psi_1 \vee \psi_2 \in \mathcal{B}_+$, that $\psi_i \prec \psi_1 \vee \psi_2$ for $i = 1,2$; and that if $\psi \in \mathcal{B}_+$ with $\psi_i \prec \psi$ for $i = 1,2$, then $\psi_1 \vee \psi_2 \prec \psi$.

We show that $\psi_1 \wedge \psi_2 \in \mathcal{B}_+$. First, obtain that $t \leq \psi_i(t)$ for all $t \in \mathbb{R}$, for $i = 1,2$, from (1.2)(i) and (ii); and then use this to obtain that $\psi_1 \wedge \psi_2$ is real-valued and that $\lim_{t\uparrow\infty}((\psi_1 \wedge \psi_2)(t) - t)0$. By definition, $\psi_1 \wedge \psi_2$ is convex. Also, $\psi_1 \wedge \psi_2$ is nondecreasing; if this were not true, then one could find $x < y$ with $\psi_1(x) \wedge \psi_2(x) > \psi_1(y) \wedge \psi_2(y)$, which is impossible since ψ_1 and ψ_2 are both nondecreasing (e.g., use ([13]: page 37)). If we let D_+u denote the right-hand derivative of function u, we have $D_+(\psi_1 \wedge \psi_2) \geq 0$, and $D_+(\psi_1 \wedge \psi_2)(t) \downarrow$ as $t \downarrow -\infty$. To show $\psi_1 \wedge \psi_2$ satisfies (1.2)(iii), it suffices to show

$$\lim_{t\downarrow-\infty} D_+(\psi_1 \wedge \psi_2)(t) = 0.$$

But we have for each $t \in \mathbb{R}$, there is some $\bar{t} \leq t$ such that

$$D_+(\psi_1 \wedge \psi_2)(t) \leq D_+(\psi_1)(\bar{t}) \vee D_+(\psi_2)(\bar{t})$$

and so the result follows since $\lim_{t\downarrow-\infty} D_+(\psi_1)(t) \vee D_+(\psi_2)(t) = 0$. This shows $\psi_1 \wedge \psi_2 \in \mathcal{B}_+$. From the definition of $\psi_1 \wedge \psi_2$, it is immediate that $\psi_1 \wedge \psi_2 \prec \psi_i$ for $i = 1,2$; and that if $\psi \in \mathcal{B}_+$ with $\psi \prec \psi_i$ for $i = 1,2$, then $\psi \prec \psi_1 \wedge \psi_2$. ∎

Now, define the mapping Γ from $\mathcal{E}_+$ into $\mathcal{B}_+$ by

$$(\Gamma(\mu))(t) = \psi_\mu(t) := \int (x \vee t)\, d\mu(x), \qquad \text{for } \mu \in \mathcal{E}_+.$$

One can use the representations

$$\psi_\mu(t) - t = \int (x-t)_+ d\mu(x) = \int_t^\infty (x-t)\, d\mu(x) = \int_t^\infty (1 - F(x))\, dx$$

to see that ψ_μ is an element of $\mathcal{B}_+$. Through the mapping Γ, the lattice structure of $(\mathcal{E}_+, \prec_c)$ becomes apparent.

PROPOSITION 1.4: *The mapping Γ is an order-preserving isomorphism from $(\mathcal{E}_+, \prec_c)$ to $(\mathcal{B}_+, \prec)$. The space $(\mathcal{E}_+, \prec_c)$ is a lattice under the operations $\vee_c$ and $\wedge_c$, defined for μ_1 and μ_2 in $\mathcal{E}_+$ by*

$$\mu_1 \vee_c \mu_2 = \Gamma^{-1}(\psi_{\mu_1} \vee \psi_{\mu_2}) \quad \text{and} \quad \mu_1 \wedge_c \mu_2 = \Gamma^{-1}(\psi_{\mu_1} \wedge \psi_{\mu_2}). \tag{1.4}$$

Proof: First, observe from the definitions that for $\mu_1, \mu_2 \in \mathcal{E}_+, \mu_1 \prec_c \mu_2$ if and only if $\psi_{\mu_1}(t) \le \psi_{\mu_2}(t)$ for all $t \in \mathbb{R}$. The one-to-one property of Γ follows from the anti-symmetry of the partial order $\prec_c$. The onto property of Γ is shown by taking any function satisfying (1.2); defining $F(x) := D_+\psi(x)$, the right-hand derivative of ψ at x; and showing that F is a distribution function with $\int_0^\infty x dF(x) < \infty$ from the properties (1.2). Thus, the inverse mapping Γ^{-1} from $\mathcal{B}_+$ into $\mathcal{E}_+$ is identified by $(\Gamma^{-1}(\psi))((-\infty, x]) = D_+\psi(x)$ for $\psi \in \mathcal{B}_+$. Now, use Lemma 1.3 to show that $(\mathcal{E}_+, \prec_c)$ is a lattice, and that the mapping Γ is order-preserving. ∎

Third, define the order $\prec_k$ on the set of p.m.'s $\mathcal{E}$ of (1.1) by $\mu_1 \prec_k \mu_2$ if $\int \phi d\mu_1 \le \int \phi d\mu_2$ for all convex functions ϕ for which both integrals exist. Now, $\prec_k$ is a partial order on $\mathcal{E}$, but the set $\mathcal{E}$ is not a lattice in the $\prec_k$ order. To see this first observe that if $\mu_1, \mu_2 \in \mathcal{E}$ with $\mu_1 \prec_k \mu_2$, then $\int x d\mu_1 = \int x d\mu_2$. Thus, for any two elements $\nu_1, \nu_2 \in \mathcal{E}$ with $\int x d\nu_1 \ne \int x d\nu_2$, there is no element $\mu \in \mathcal{E}$ with $\mu \prec_k \nu_1$ and $\mu \prec_k \nu_2$. However, this suggests that we consider, for each $r \in \mathbb{R}$, the set

$$\mathcal{E}_r = \{\mu \in \mathcal{P}(\mathbb{R}) : \int x d\mu(x) = r\}.$$

We will show that $(\mathcal{E}_r, \prec_k)$ is a lattice by using the following.

LEMMA 1.5:

(i) *$(\mathcal{E}, \prec_c)$ is a lattice under the operations $\vee_c$ and $\wedge_c$ of (1.4).*

(ii) *Let μ_1 and μ_2 be in $\mathcal{E}$. Then $\mu_1 \prec_k \mu_2$ if and only if $\mu_1 \prec_c \mu_2$ and $\int x d\mu_1 = \int x d\mu_2$.*

Proof: To see (i), replace the set $\mathcal{B}_+$ by the set $\mathcal{B}$ of functions ψ on $\mathbb{R}$ satisfying (1.2)(i), (ii) and (iii)′ ψ is bounded from below. Then use reasoning analogous to that leading to Proposition 1.4. In particular, one shows that $(\mathcal{B}, \prec)$ is a lattice under the operations of (1.3); and one uses the observation that for $\mu \in \mathcal{E}, \lim_{t \downarrow -\infty} \psi_\mu(t) = \int x d\mu(x)$, to show that the mapping Γ is an order-preserving isomorphism from $(\mathcal{E}, \prec_c)$ to $(\mathcal{B}, \prec)$, and that the space $(\mathcal{E}, \prec_c)$ is a lattice under the operations of (1.4). For conclusion (ii), see ([14]: page 9). ∎

The result that the mapping Γ is a one-to-one mapping from $\mathcal{E}$ onto $\mathcal{B}$ was given by Gilat [7].

PROPOSITION 1.6: *Fix any $r \in \mathbb{R}$. The space $(\mathcal{E}_r, \prec_k)$ is a lattice under operations $\vee_k$ and $\wedge_k$ defined to be the same as those under the $\prec_c$ order, given in (1.4).*

Proof: Let $\mu_1, \mu_2 \in \mathcal{E}_r$. Now, $\mu_1 \wedge_c \mu_2$ and $\mu_1 \vee_c \mu_2$ are in $\mathcal{E}$. We show

$$\int x d(\mu_1 \wedge_c \mu_2) = r = \int x d(\mu_1 \vee_c \mu_2).$$

We have $\int x d(\mu_1 \wedge_c \mu_2) = r$ since $\varepsilon_r \prec_c \mu_1 \wedge_c \mu_2 \prec_c \mu_i$ for $i = 1, 2$ (where ε_r denotes point mass at r); and we have $\int x d(\mu_1 \vee_c \mu_2) = r$ since

$$\int x d(\mu_1 \vee_c \mu_2) = \lim_{t \downarrow -\infty} \psi_{\mu_1 \vee_c \mu_2}(t) = \lim_{t \downarrow -\infty} \psi_{\mu_1}(t) \vee \psi_{\mu_2}(t) = r.$$

Thus $\mu_1 \wedge_c \mu_2$ and $\mu_1 \vee_c \mu_2$ are in $\mathcal{E}_r$, and $\mu_1 \wedge_c \mu_2 \prec_k \mu_i \prec_k \mu_1 \vee_c \mu_2$ for $i = 1, 2$ from Lemma 1.5(ii). The remainder of the proof is straightforward. For example, if $\mu \in \mathcal{E}_r$ and $\mu \prec_k \mu_i$ for $i = 1, 2$, then $\mu \prec_c \mu_i$ for $i = 1, 2$ and so $\mu \prec_c \mu_1 \wedge_c \mu_2$, thus also $\mu \prec_k \mu_1 \wedge_c \mu_2$. ∎

For p.m.'s in $\mathcal{E}$, there is a known connection between the orders $\prec_c$ and $\prec_k$ and processes of submartingales and martingales. From Theorems 8 and 9 of [15], one has the following.

LEMMA 1.7: *Let μ_1 and μ_2 be p.m.'s in $\mathcal{E}$.*

(a) *Each of the following is equivalent to $\mu_1 \prec_c \mu_2$:*

(i) *(μ_1, μ_2) is a submartingale pair, that is, there is a probability space $(\Omega, \mathcal{F}, P)$, and r.v.'s X_1 and X_2 on Ω with associated p.m.'s μ_1 and μ_2 respectively, for which $X_1 \le E(X_2|X_1)$ a.e. [P];*

(ii) *there is a probability kernel $K : \mathbb{R} \times \mathcal{B}(\mathbb{R}) \to [0, 1]$ with $x \le \int y K(x, dy)$ for all $x \in \mathbb{R}$ and $\int K(x, A)\mu_1(dx) = \mu_2(A)$ for all $A \in \mathcal{B}(\mathbb{R})$.*

(b) *Each of the following is equivalent to $\mu_1 \prec_k \mu_2$:*

(i) *(μ_1, μ_2) is a martingale pair, i.e., there is a probability space $(\Omega, \mathcal{F}, P)$, and r.v.'s X_1 and X_2 on Ω with associated p.m.'s μ_1 and μ_2 respectively, for which $X_1 = E(X_2|X_1)$ a.e. [P];*

(ii) *there is a probability kernel $K : \mathbb{R} \times \mathcal{B}(\mathbb{R}) \to [0, 1]$ with $\int y K(x, dy) = x$ for all $x \in \mathbb{R}$ and $\int K(x, A)\mu_1(dx) = \mu_2(A)$ for all $A \in \mathcal{B}(\mathbb{R})$.*

In this process context, Lemma 1.5 and Proposition 1.6 have the following interpretations. Let μ_1 and μ_2 be in $\mathcal{E}$. Then $\mu_1 \wedge_c \mu_2$ is the unique p.m. on $\mathbb{R}$

satisfying (i) $(\mu_1 \wedge_c \mu_2, \mu_1)$ and $(\mu_1 \wedge_c \mu_2, \mu_2)$ are both submartingale pairs; and (ii) if μ is any element of $\mathcal{E}$ for which both (μ, μ_1) and (μ, μ_2) are submartingale pairs, then $(\mu, \mu_1 \wedge_c \mu_2)$ is also a submartingale pair. Next, let μ_1 and μ_2 be in $\mathcal{E}_r$, for fixed $r \in \mathbb{R}$. Then $\mu_1 \wedge_k \mu_2$ is the unique p.m. on $\mathbb{R}$ satisfying (i) $(\mu_1 \wedge_k \mu_2, \mu_1)$ and $(\mu_1 \wedge_k \mu_2, \mu_2)$ are both martingale pairs; and (ii) if μ is any element of $\mathcal{E}_r$ for which both (μ, μ_1) and (μ, μ_2) are martingale pairs, then $(\mu, \mu_1 \wedge_k \mu_2)$ is also a martingale pair. Analogous interpretations can be given to $\mu_1 \vee_c \mu_2$ and $\mu_1 \vee_k \mu_2$.

In this paper, we are mainly concerned with the $\prec_s$ and $\prec_c$ orders. There are some useful relationships between these orders. It is immediate that for $\mu_1, \mu_2 \in \mathcal{E}_+$, if $\mu_1 \prec_s \mu_2$, then $\mu_1 \prec_c \mu_2$. We develop a more precise connection between the orders $\prec_s$ and $\prec_c$. For p.m. $\mu \in \mathcal{E}_+$, let $H^{-1} = H_\mu^{-1}$ denote the Hardy and Littlewood maximal function associated with μ, defined on [0,1] by

$$H^{-1}(w) = (1-w)^{-1} \int_w^1 F_\mu^{-1}(u)du$$

(with $H^{-1}(1) = F_\mu^{-1}(1)$), having associated distribution function $H = H_\mu$ (with left inverse H^{-1}) and associated p.m. μ^*, the Hardy and Littlewood maximal p.m. associated with μ. The following lemma is basic to our analysis; it equates the $\prec_c$ order on p.m.'s and the $\prec_s$ order on their maximal p.m.'s.

LEMMA 1.8: *For p.m.'s μ_1 and μ_2 in $\mathcal{E}_+, \mu_1 \prec_c \mu_2$ if and only if $\mu_1^* \prec_s \mu_2^*$.*

Proof: Let μ_1 and μ_2 be p.m.'s in $\mathcal{E}_+$ having distribution functions F_1 and F_2 respectively. Define the crossover points associated with F_1 and F_2 as follows: (c, w) is a crossover point if for $i \neq j \in 1, 2$,

$$F_i(c-) \le F_j(c-) \le F_j(c) \le F_i(c) \quad \text{and} \quad w = F_j(c),$$

or

$$F_i^{-1}(w) \le F_j^{-1}(w) \le F_j^{-1}(w+) \le F_i^{-1}(w+) \quad \text{and} \quad c = F_j^{-1}(w).$$

We have

$$\begin{aligned}\mu_1^* \prec_s \mu_2^* &\Leftrightarrow H_{\mu_1}^{-1}(w) \le H_{\mu_2}^{-1}(w) \quad \text{for all } w \in [0,1]\\ &\Leftrightarrow \int_w^1 (F_2^{-1}(u) - F_1^{-1}(u))du \geq 0 \quad \text{for all crossover points } (c,w)\\ &\Leftrightarrow \int_c^\infty (F_1(u) - F_2(x))dx \ge 0 \quad \text{for all crossover points } (c,w)\\ &\Leftrightarrow \int_c^\infty (1 - F_1(x))dx \le \int_c^\infty (1 - F_2(x))dx \quad \text{for all crossover points } (c,w)\\ &\Leftrightarrow \mu_1 \prec_c \mu_2,\end{aligned}$$

where we have used that for each crossover point (c, w),

$$\begin{aligned}\int_w^1 (F_2^{-1}(u) - F_1^{-1}(u))du &= \int_c^\infty (F_1(x) - F_2(x))dx\\ &= \text{area between } F_1 \text{ and } F_2 \text{ to the right of } c\\ &= \text{area between } F_1^{-1} \text{ and } F_2^{-1} \text{ to the right of } w. \ \blacksquare\end{aligned}$$

The following is now immediate:

$$\text{for } \mu_1 \text{ and } \mu_2 \text{ in } \mathcal{E} \text{ with the same mean,}$$
$$\mu_1 \prec_c \mu_2 \Leftrightarrow \mu_1 \prec_k \mu_2 \Leftrightarrow \mu_1^* \prec_s \mu_2^*. \tag{1.5}$$

Under the common mean assumption, van der Vecht ([16]: page 69) gave a result (attributed to D. Gilat) equivalent to the second equivalence in (1.5).

2. Existence and Examples of Minimal Probability Measures

Let ν be any p.m. on $\mathbb{R}$, and recall from (1.1) that $\mathcal{E}_+$ is the set of p.m.'s μ on $\mathbb{R}$ satisfying $\int_0^\infty x d\mu < \infty$. Let $\mathcal{E}_+(\nu)$ be the subset of $\mathcal{E}_+$ given by

$$\mathcal{E}_+(\nu) = \{\mu \in \mathcal{E}_+ : \nu \prec_s \mu^*\}. \tag{2.1}$$

Theorem A of the Introduction states, in part, that if $\mathcal{E}_+(\nu)$ is nonempty, then it contains its greatest lower bound in the $\prec_c$ order, which is called ν_Δ. Verification of Theorem A follows easily from Proposition 2.1 and Theorem 2.4. We call this p.m. ν_Δ the **minimal probability measure** associated with ν. Examples of minimal p.m.'s are given after Theorem 2.4.

The set of p.m.'s ν which have a minimal p.m. ν_Δ is the set of p.m.'s ν for which $\mathcal{E}_+(\nu)$ is nonempty; this set is denoted $\mathcal{P}^*$, that is,

$$\mathcal{P}^* = \{\nu \in \mathcal{P}(\mathbb{R}) : \nu \prec_s \mu^* \text{ for some p.m. } \mu \text{ with } \int_0^\infty x d\mu < \infty\}. \tag{2.2}$$

We characterize the set $\mathcal{P}^*$ in the following.

PROPOSITION 2.1: $\mathcal{P}^* = \{\nu \in \mathcal{P}(\mathbb{R}) : \limsup_{x\to\infty} x\nu[x,\infty) = 0\}$.

Proof: We first prove the containment '$\subset$'. Let $\nu \in \mathcal{P}^*$. We suppose $\nu[x,\infty) > 0$ for all $x \in \mathbb{R}$; otherwise the containment is immediate. Let $\mu \in \mathcal{E}_+$ such that $\nu \prec_s \mu^*$. Now, $H = H_\mu = F_{\mu^*}$ is strictly increasing and continuous, and

$$x\mu^*[x,\infty) = x(1-H(x)) = H^{-1}(H(x))(1-H(x)) = \int_{H(x)}^1 F_\mu^{-1}(t)dt.$$

From the assumptions, it follows directly that

$$0 \le \limsup_{x\uparrow\infty} x\nu[x,\infty) \le \limsup_{x\uparrow\infty} x\mu^*[x,\infty) = \limsup_{x\uparrow\infty} \int_{H(x)}^1 F_\mu^{-1}(t)dt = 0$$

and the containment is proved.

Next, we prove the containment '$\supset$'. Let ν be any p.m. on $\mathbb{R}$ satisfying

$$\limsup_{x\uparrow\infty} x\nu[x,\infty) = 0. \tag{2.3}$$

We suppose that ν is a discrete p.m. on $\mathbb{R}$ having atoms $(y_n)_{n=0,1,\ldots}$ such that $0 < y_0 < \ldots < y_n < y_{n+1} < \cdots \uparrow \infty$ and (y_n) has no accumulation point. (If ν is any p.m. satisfying (2.3), then one can always construct a discrete p.m. $\bar\nu$ such that $\limsup_{x\uparrow\infty} x\bar\nu[x,\infty) = 0$, the atoms of $\bar\nu$ have the desired properties, and $\nu \prec_s \bar\nu$; and then one may reduce the argument to working with $\bar\nu$.) We construct a p.m. μ on $\mathbb{R}$ satisfying $\int_0^\infty y d\mu < \infty$ and $\nu \prec_s \mu^*$.

Define the function $\Lambda(x) = \Lambda(x;\nu)$ on (y_n) by

$$\Lambda(x) = \inf_{z>x}(z\nu[z,\infty) - x\nu[x,\infty))/(\nu[z,\infty) - \nu[x,\infty)).$$

One uses the properties of ν in a straightforward way to show that for each atom x in (y_n),

(i) $-\infty < \Lambda(x) < x$,

(ii) $(z\nu[z,\infty) - x\nu[x,\infty))/(\nu[z,\infty) - \nu[x,\infty)) \to x$ as $z \downarrow x$ and as $z \uparrow \infty$,

(iii) $\Lambda(x) = (y\nu[y,\infty) - x\nu[x,\infty))/(\nu[y,\infty) - \nu[x,\infty))$ for some $y \in (y_n)$.

Thus, there is a subsequence (x_n) of (y_n) with $x_0 < \cdots < x_n < x_{n+1} < \cdots$ constructed recursively as follows: let $x_0 = y_0$, and if $x_1, \ldots, x_{j-1}$ have been found, then define x_j as the largest number y in (y_n) for which

$$\Lambda(x_{j-1}) = (y\nu[y,\infty) - x_{j-1}\nu[x_{j-1},\infty))/(\nu[y,\infty) - \nu[x_{j-1},\infty)).$$

Let a_n, b_n, c_n, for $n = 0, 1, \ldots$, be parameters defined by

$$\text{(2.4)} \qquad \begin{aligned} a_n &= \nu[x_n,\infty); \qquad b_n = \nu[x_n, x_{n+1}) = a_n - a_{n+1}; \qquad \text{and} \\ c_n &= \Lambda(x_n) = (x_{n+1}a_{n+1} - x_n a_n)/(a_{n+1} - a_n); \end{aligned}$$

and define p.m. μ by $\mu = \Sigma_{n=0}^{\infty} b_n \varepsilon_{c_n}$. First observe that $\mu \in \mathcal{E}_+$, since

$$\begin{aligned} \int y d\mu &= \Sigma_{n=0}^{\infty} b_n c_n = \Sigma_{n=0}^{\infty}(x_n a_n - x_{n+1}a_{n+1}) \\ &= x_0 - \lim_n a_n x_n = x_0 - \lim_n x_n \nu[x_n,\infty) = x_0 \end{aligned}$$

from (2.3) and (2.4).

Next, we show that $\nu \prec_s \mu^*$ by verifying that $\nu[x,\infty) \le \mu^*[x,\infty)$ for all $x \in \mathbb{R}$. One obtains through direct calculation (e.g., with use of F_μ, F_μ^{-1}, and H_μ^{-1}) that for $x_{j-1} \le x \le x_j$,

$$\text{(2.5)} \qquad \begin{aligned} \mu^*[x,\infty) &= \nu[x_j,\infty)(x_j - c_{j-1})/(x - c_{j-1}) \\ &= \nu[x_{j-1},\infty)(x_{j-1} - c_{j-1})/(x - c_{j-1}). \end{aligned}$$

Thus, for $x_{j-1} \le x \le x_j$, $\nu[x,\infty) \le \mu^*[x,\infty)$ is equivalent to

$$\text{(2.6)} \qquad \nu[x,\infty) \le \nu[x_{j-1},\infty)(x_{j-1} - c_{j-1})/(x - c_{j-1}).$$

But, by algebraic manipulation one obtains that (2.6) is equivalent to

$$c_{j-1} \le (x\nu[x,\infty) - x_{j-1}\nu[x_{j-1},\infty))/(\nu[x,\infty) - \nu(x_{j-1},\infty)),$$

and this is immediate from the definition of $c_{j-1} = \Lambda(x_{j-1})$. ∎

Examples of probability measures in $\mathcal{P}^*$ are the following:
(i) $\mathcal{E}_+ \subset \mathcal{P}^*$ (e.g., since if $\nu \in \mathcal{E}_+$, then $\nu \in \mathcal{E}_+(\nu)$), for example, p.m.'s with bounded support; Gaussian, exponential, and gamma p.m.'s, etc.
(ii) $\mathcal{P}_0^* \subset \mathcal{P}^*$ (e.g., since if $\nu = \mu^*$ for some $\mu \in \mathcal{E}_+$, then $\mu \in \mathcal{E}_+(\nu)$), where

$$\mathcal{P}_0^* := \left\{ \nu \in \mathcal{P}(\mathbb{R}) : \nu = \mu^* \text{ for some p.m. } \mu \text{ with } \int_0^\infty y d\mu < \infty \right\}, \tag{2.7}$$

the set of maximal p.m.'s.

Note that there are p.m.'s in $\mathcal{E}_+ \setminus \mathcal{P}_0^*$ (e.g., any nondegenerate p.m. ν concentrated on finitely many atoms); and that there are p.m.'s in $\mathcal{P}_0^* \setminus \mathcal{E}_+$ (in particular, if $\int_0^\infty x d\mu < \infty$ but $\int_0^\infty x\,(\ln x)_+ d\mu = \infty$, then it may follow that $\int_0^\infty x d\mu^* = \infty$; see, e.g., [11]). Examples of probability measures not in $\mathcal{P}^*$ are those in the domain of attraction (for maxima) of the second extremal type distribution function $\Phi_\alpha(x) = \exp\{-x^{-\alpha}\}, x > 0$, for $0 < \alpha \leq 1$, for example, the Pareto distribution $0 < \alpha \leq 1$ and the Cauchy distribution; use the representation for distribution functions in this class given, e.g., in ([12]: Corollary 1.12).

We develop some lemmas that are used in the proof of Theorem 2.4. In the following, we use $\mathbb{N} = \{1, 2, \ldots\}$ and use $\rho_n \Rightarrow \rho$ to denote weak convergence of p.m.'s $\{\rho_n\}$ to p.m. ρ (convergence in distribution of $\{F_{\rho_n}\}$ to F_ρ). We use the following equivalence (see, e.g., page 5 of [12]) for p.m.'s $\{\rho_n\}$ and ρ:

$$\begin{aligned} &\rho_n \Rightarrow \rho \text{ if and only if } \lim_n F_{\rho_n}^{-1}(t) = F_\rho^{-1}(t) \\ &\text{for each continuity point } t \text{ of } F_\rho^{-1} \text{ in } (0,1). \end{aligned} \tag{2.8}$$

Lemma 2.2: *Let $\nu \in \mathcal{P}(\mathbb{R})$ and $\mu_0 \in \mathcal{E}_+$ with $\nu \prec_s \mu_0^*$. Then the set of p.m.'s $\{\rho \in \mathcal{P}_0^* : \nu \prec_s \rho \prec_s \mu_0^*\}$ is sequentially compact in the topology of weak convergence.*

Proof: Let $\{\mu_n\}_{n \in \mathbb{N}}$ be any sequence of p.m.'s in $\mathcal{E}_+$ satisfying $\nu \prec_s \mu_n^* \prec_s \mu_0^*$ for all $n \in \mathbb{N}$. For each $n = 0, 1, 2, \ldots$, associate with p.m. μ_n the distribution function F_n and left-inverse function F_n^{-1}, and maximal p.m. μ_n^* together with its distribution function H_n and left-inverse function H_n^{-1}. There is a subsequence $\{i\}$ of $\mathbb{N}$ with $F_i(u)$ converging in i to some value, say $F(u)$, for all u in some dense subset D of $\mathbb{R}$. Then $F : D \to [0,1]$ is nondecreasing and can uniquely be extended as a right-continuous function to $F : \mathbb{R} \to [0,1]$.

We show F is a distribution function. First, we show that $\lim_{t\uparrow\infty} F(t) = 1$. Suppose this is not true; then there is some $t_1 \in (0,1)$ at which $F_i^{-1}(t_1)$ converges to ∞. Since $\mu_n^* \prec_s \mu_0^*$ for all $n \in \mathbb{N}$, we obtain the contradiction

$$\infty = \lim_i (1-t_1)^{-1} \int_{t_1}^{1} F_i^{-1}(u)du \leq (1-t_1)^{-1} \int_{t_1}^{1} F_0^{-1}(u)du = H_0^{-1}(t_1) < \infty.$$

Next, we show that $\lim_{t\downarrow-\infty} F(t) = 0$. Suppose this is not true; then there is some $s_1 \in (0,1)$ at which $F_i^{-1}(s_1)$ converges to $-\infty$. But then again using $\mu_n^* \prec_s \mu_0^*$ for all $n \in \mathbb{N}$, we obtain for each $s_0 \in (0, s_1)$,

$$\int_{s_0}^{1} F_i^{-1}(u)du \leq \int_{s_0}^{s_1} F_i^{-1}(u)du + (1-s_1)H_0^{-1}(s_1),$$

and therefore

$$\liminf_i H_i^{-1}(s_0) = -\infty \qquad \text{for each } s_0 \in (0, s_1). \tag{2.9}$$

But we know $\nu \prec_s \mu_n^*$ for all $n \in \mathbb{N}$; and so for all $x \in \mathbb{R}, \nu(-\infty, x] \geq \mu_i^*(-\infty, x] = H_i(x)$, and $H_i^{-1}(\nu(-\infty,x]) \geq x$. But this, together with (2.9), forces $\nu(-\infty, x] \geq s_1$ for all $x \in \mathbb{R}$, a contradiction to ν being a p.m. on $\mathbb{R}$. Thus F is a distribution function.

Associated with the distribution function F are the p.m. μ and left-inverse function F^{-1}. We have $\mu_i \Rightarrow \mu$. We define

$$H^{-1}(t) = (1-t)^{-1} \int_{t}^{1} F^{-1}(u)du \qquad \text{for } t \in (0,1)$$

and will show that

$$H_i^{-1}(t) \to H^{-1}(t) \qquad \text{as } n \uparrow \infty, \qquad \text{for each } t \in (0,1). \tag{2.10}$$

From this it easily follows that $\mu \in \mathcal{E}_+$; the maximal p.m. μ^* has distribution function H and left-inverse function $H^{-1}; \nu \prec_s \mu^* \prec_s \mu_0^*$; and $\mu_i^* \Rightarrow \mu^*$ from

(2.8). For the proof of (2.10), we first obtain that

$$\text{(2.11) (i)} \qquad \lim_i \int_{t_0}^{t_1} F_i^{-1}(u)du = \int_{t_0}^{t_1} F^{-1}(u)du \text{ for all } 0 < t_0 < t_1 < 1;$$

$$\text{(ii)} \lim_{t_1 \uparrow 1} \left(\liminf_i \int_{t_1}^{1} F_i^{-1}(u)du \right) = 0 = \lim_{t \downarrow 1} \left(\limsup_i \int_{t_1}^{1} F_i^{-1}(u)du \right);$$

$$\text{(iii)} \lim_{t_1 \uparrow 1} \int_{t_1}^{1} F^{-1}(u)du = 0.$$

To verify that (2.11)(i) holds, observe that $\{F_i^{-1}(u)\}$ is uniformly bounded over i and u in $[t_0, t_1]$, use that $F_i^{-1}(u) \to F^{-1}(u)$ for each continuity point u of F^{-1} in (0,1), and apply the bounded convergence theorem. To verify that (2.11)(ii) holds, use that $\nu \prec_s \mu_n^* \prec_s \mu_0^*$ for all $n \in \mathbb{N}$ to obtain for $0 < \bar{t} < t < 1$,

$$(1-t)F_\nu^{-1}(\bar{t}) \le (1-t)F_\nu^{-1}(t) \le \liminf_n \int_t^1 F_n^{-1}(u)du$$

$$\le \limsup_n \int_t^1 F_n^{-1}(u)du \le \int_t^1 F_0^{-1}(u)du$$

and let $t \uparrow 1$ (observe $\int_t^1 F_0^{-1}(u)du \to 0$ since $\int_0^\infty x d\mu_0(x) < \infty$). To see that (2.11)(iii) holds, observe that if (2.11)(iii) fails, it must be the case that

$$\lim_{t_1 \uparrow 1} \int_{t_1}^{1} F^{-1}(u)du = +\infty.$$

But then for any $\varepsilon > 0$, for each t_0 close to 1, there is t_1 in $(t_0, 1)$ for which

$$\varepsilon < \int_{t_0}^{t_1} F^{-1}(u)du = \lim_i \int_{t_0}^{t_1} F_i^{-1}(u)du \le \liminf_i \int_{t_0}^{1} F_i^{-1}(u)du,$$

where we have used (2.11)(i). But this contradicts (2.11)(ii). Now, use (2.11)(i) to obtain for all $t_0 < t_1 < 1$

$$\liminf_i \int_{t_0}^{1} F_i^{-1}(u)du - \int_{t_0}^{1} F^{-1}(u)du = \liminf_i \int_{t_1}^{1} F_i^{-1}(u)du - \int_{t_1}^{1} F^{-1}(u)du.$$

Using this and (2.11)(ii), (iii) and letting $t_1 \uparrow 1$, we obtain

$$\begin{aligned}
&\liminf_i \int_{t_0}^{1} F_i^{-1}(u)du - \int_{t_0}^{1} F^{-1}(u)du \\
&= \lim_{t_1 \uparrow 1} \left(\liminf_i \int_{t_1}^{1} F_i^{-1}(u)du - \int_{t_1}^{1} F^{-1}(u)du \right) \\
&= 0.
\end{aligned}$$

By an entirely analogous argument, one uses (2.11) to show

$$\limsup_i \int_{t_0}^{1} F_i^{-1}(u)du = \int_{t_0}^{1} F^{-1}(u)du.$$

The convergence in (2.10) now follows. ∎

Lemma 2.3: *Let $\{\mu_n\}_{n\in\mathbb{N}}$ be any sequence of p.m.'s in $\mathcal{E}_+$ satisfying $\nu \prec_s \mu_n^* \prec_s \mu_0^*$ for all $n \in \mathbb{N}$, for some $\nu \in \mathcal{P}(\mathbb{R})$ and $\mu_0 \in \mathcal{E}_+$. Then*

(a) *if $\{\mu_n^*\}$ is decreasing in the $\prec_s$ order, then there is a unique p.m. $\mu \in \mathcal{E}_+$ with $\mu_n^* \Rightarrow \mu^*$ and $\mu^* = \wedge_s\{\mu_n^* : n \in \mathbb{N}\}$; and*

(b) *if $\{\mu_n^*\}$ is increasing in the $\prec_s$ order, then there is a unique p.m. $\mu \in \mathcal{E}_+$ with $\mu_n^* \Rightarrow \mu^*$ and $\mu^* = \vee_s\{\mu_n^* : n \in \mathbb{N}\}$.*

Proof: We prove part (a); the proof of part (b) is analogous. Let $\{\mu_n\}_{n\in\mathbb{N}}$ be p.m.'s satisfying the hypotheses under part (a), with associated maximal p.m.'s $\{\mu_n^*\}$ having left-inverse functions $\{H_n^{-1}\}$. Then $\{H_n^{-1}\}$ are pointwise decreasing in n. From Lemma 2.2, there is a p.m. $\mu \in \mathcal{E}_+$ with associated maximal p.m. μ^* having left-inverse function H^{-1} for which $\mu_n^* \Rightarrow \mu^*$ and $H_n^{-1}(t) \downarrow H^{-1}(t)$ for each $t \in (0,1)$ as $n \to \infty$. Since $H^{-1} \le H_n^{-1}$, we have $\mu^* \prec_s \mu_n^*$, for all $n \in \mathbb{N}$. Also, if $\rho \in \mathcal{P}(\mathbb{R})$ with $\rho \prec_s \mu_n^*$ for all $n \in \mathbb{N}$, then $F_\rho^{-1} \le H_n^{-1}$ for all $n \in \mathbb{N}$, and one concludes from the convergence that $F_\rho^{-1} \le H^{-1}$ and $\rho \prec_s \mu^*$. Thus $\mu^* = \wedge_s\{\mu_n^* : n \in \mathbb{R}\}$. ∎

Under either the assumptions of part (a) or of part (b) of Lemma 2.2, it follows easily (using ideas in the proof of Lemma 2.2 for (2.12) (ii)) that the p.m. μ in the conclusion satisfies

(2.12) (i) $\nu \prec_s \mu^* \prec_s \mu_0^*$, and

(ii) if p.m. λ satisfies $\mu_n \prec_s \lambda$ for all $n \in \mathbb{R}$, then $\mu \prec_s \lambda$.

THEOREM 2.4: *Let ν be any p.m. on $\mathbb{R}$ satisfying $\nu \prec_s \mu_0^*$ for some p.m. μ_0 with $\int_0^\infty x d\mu_0 < \infty$. Then there is a unique p.m. ν_Δ on $\mathbb{R}$ satisfying* (i) $\int_0^\infty x d\nu_\Delta(x) < \infty$; (ii) $\nu \prec_s (\nu_\Delta)^*$; *and* (iii) *if $\bar{\mu}$ is any p.m. on $\mathbb{R}$ with $\int_0^\infty x d\bar{\mu}(x) < \infty$ and $\nu \prec_s \bar{\mu}^*$, then $(\nu_\Delta)^* \prec_s \bar{\mu}^*$.*

Proof: Let D denote a countable dense subset of $\mathbb{R}$. From the usual diagonalization procedure, there is a sequence $\{\mu_n : n = 1, 2, \ldots\}$ in $\mathcal{E}_+(\nu)$ satisfying, for all $x \in D$,

$$\sup_{n \in \mathbb{N}} \mu_n^*(-\infty, x] = \sup_{\rho \in \mathcal{E}_+(\nu)} \rho^*(-\infty, x]. \tag{2.13}$$

Furthermore, from the lattice structure of $(\mathcal{E}_+, \prec_c)$, we may assume that $\{\mu_n\}$ is decreasing in the $\prec_c$ order and $\mu_1 \prec_c \mu_0$, so that $\{\mu_n^*\}$ is decreasing in the $\prec_s$ order and $\mu_1^* \prec_s \mu_0^*$. The hypotheses of Lemma 2.3(a) are satisfied for this sequence $\{\mu_n\}$, and we obtain unique p.m. $\mu \in \mathcal{E}_+$ with $\mu^* = \wedge_s\{\mu_n^* : n \in \mathbb{N}\}$. Define $\nu_\Delta := \mu$ and observe that conclusions (i) and (ii) are immediate from Lemma 2.3 and (2.12). To verify conclusion (iii), let $\bar{\mu}$ be any p.m. on $\mathbb{R}$ with $\int_0^\infty x d\bar{\mu} < \infty$ and $\nu \prec_s \bar{\mu}^*$. Use that $\nu_\Delta = \mu \in \mathcal{E}_+(\nu)$ (by conclusions (i) and (ii)) and (2.13) to obtain that $(\nu_\Delta)^*(-\infty, x] \geq \bar{\mu}^*(-\infty, x]$ for all $x \in D$ and thus $(\nu_\Delta)^* \prec_s \bar{\mu}^*$. Uniqueness of the p.m. ν_Δ in the conclusion is clear. ∎

The proof of Theorem A is immediate from Theorem 2.4. As a consequence of Theorem A and the equivalences of Section 1 we have that for each p.m. $\nu \in \mathcal{P}^*$, the minimal p.m. ν_Δ is in $\mathcal{E}_+(\nu)$ and satisfies

(2.14) (i) $$\int \phi d\nu_\Delta = \min_{\mu \in \mathcal{E}_+(\nu)} \int \phi d\mu$$

for every nondecreasing convex function for which the integrals exist, and

(ii) $$\int_t^\infty \nu_\Delta[x, \infty) dx = \min_{\mu \in \mathcal{E}_+(\nu)} \int_t^\infty \mu[x, \infty) dx \quad \text{for all } t \in \mathbb{R}$$

and the associated maximal p.m. $(\nu_\Delta)^*$ and its left-inverse function H_Δ^{-1} satisfy

(2.15) (i) $$(\nu_\Delta)^*(-\infty, x] = \max_{\mu \in \mathcal{E}_+(\nu)} \mu^*(-\infty, x] \quad \text{for all } x \in \mathbb{R}, \quad \text{and}$$

(ii) $$H_\Delta^{-1}(t) = \inf_{\mu \in \mathcal{E}_+(\nu)} H_\mu^{-1}(t) \quad \text{for all} \quad t \in [0, 1].$$

We now identify minimal p.m.'s ν_Δ in the special cases of maximal p.m.'s ν (Lemma 2.5) and simple discrete p.m.'s ν (Proposition 2.8), and give specific examples.

LEMMA 2.5: (a) *If $\mu \in \mathcal{E}_+$, then $(\mu^*)_\Delta = \mu$.*
(b) *If $\nu \in \mathcal{P}_0^*$, then the left-inverse function $F_{\nu_\Delta}^{-1}$ for p.m. ν_Δ is given through the equation*

$$F_{\nu_\Delta}^{-1}(w) = F_\nu^{-1}(w) - (1-w)\frac{d}{dw}F_\nu^{-1}(w). \tag{2.16}$$

Proof: For part (a), use the defining properties of $(\mu^*)_\Delta$ as given in Theorem 2.4 to obtain $\mu^* \prec_s ((\mu^*)_\Delta)^*$ and $((\mu^*)_\Delta)^* \prec_s \mu^*$; and then use Lemma 1.8. Part (a) gives that if $\nu = \mu^*$ for $\mu \in \mathcal{E}_+$, then $\nu_\Delta = \mu$; thus ν and ν_Δ are related by

$$F_\nu^{-1}(w) = (1-w)^{-1}\int_w^1 F_{\nu_\Delta}^{-1}(u)du$$

and (2.16) follows by differentiation. If necessary, one takes a left-continuous modification of the expression identified in (2.16). ∎

Part (b) of Lemma 2.5 states that, for any maximal p.m. ν, the left-inverse function $F_{\nu_\Delta}^{-1}$ for p.m. ν_Δ is a solution G of

$$F_\nu^{-1}(w) = (1-w)^{-1}\int_w^1 G(u)du,$$

and is unique a.e. among nondecreasing, left-continuous solutions G. For the purpose of constructing examples, the following characterization of the set $\mathcal{P}_0^*$ of maximal p.m.'s is useful.

LEMMA 2.6: *For any p.m. ν on $\mathbb{R}, \nu \in \mathcal{P}_0^*$ if and only if the following holds:*

$$\text{(2.17)} \qquad \text{(i)} \quad \limsup_{w\uparrow 1}(1-w)F_\nu^{-1}(w) = 0, \quad \text{and}$$

$$\text{(ii)} \quad (1-w)F_\nu^{-1}(w) \quad \text{is a concave function.}$$

Proof: If $\nu \in \mathcal{P}_0^*$, so that $\nu = \mu^*$ for some $\mu \in \mathcal{E}_+$, then (2.17)(i) is immediate and (2.17)(ii) follows from an argument similar to that in ([13]: pages 230–231), since

$$(1-w)F_\nu^{-1}(w) = \int_w^1 F_\mu^{-1}(t)dt.$$

On the other hand, if ν satisfies (2.17)(i) and (ii), let

$$a(w) = -(1-w)F_\nu^{-1}(w) \quad \text{for } w \in [0,1].$$

Since $a(w)$ is a convex function, its left derivative $a^-(w)$ is a finite, nondecreasing, left-continuous function on (0,1); extend $a^-(w)$ to [0,1] by setting $a^-(0) = a^-(0+)$ and $a^-(1) = a^-(1-)$. Consider $a^-(w)$ as a random variable on $([0,1], \mathcal{B}, m)$, and associate with $a^-(w)$ its p.m. μ and d.f. F_μ. Then $F_\mu^{-1} = a^-$ and for $0 < w < z < 1$,

$$\int_w^z F_\mu^{-1}(t)dt = \int_w^z a^-(t)dt = a(z) - a(w+) = -(1-z)F_\nu^{-1}(z) + (1-w)F_\nu^{-1}(w+).$$

Now, let $z \uparrow 1$ and use (2.17)(i) to obtain

$$F_\nu^{-1}(w+) = (1-w)^{-1} \int_w^1 F_\mu^{-1}(t)dt,$$

and thus $\nu = \mu^*$, with $\mu \in \mathcal{E}_+$. ∎

Examples 2.7: For each of the following p.m.'s ν, conditions (2.17)(i) and (ii) hold, and so $\nu \in \mathcal{P}_0^*$ and ν_Δ can be identified from (2.16).

(i) If ν is Exponential distributed with parameter one, then ν_Δ has d.f.

$$F_{\nu_\Delta}(x) = 1 - e^{-1}e^{-x} \quad \text{if } x \geq -1, \quad \text{and } = 0 \quad \text{otherwise.}$$

(ii) If ν is Pareto distributed with parameter $\alpha > 1$, so that $F_\nu(x) = 1 - x^{-\alpha}$ if $x \geq 1$, and $= 0$ otherwise, then ν_Δ has d.f. $F_{\nu_\Delta}(x) = 1 - (1-\alpha^{-1})^\alpha x^{-\alpha}$ if $x \geq 1 - \alpha^{-1}$, and $= 0$ otherwise.

(iii) If ν has d.f. $F_\nu(x) = 1 - (1-x)^\alpha$ if $0 \leq x \leq 1$, for $\alpha > 0$, then ν_Δ has d.f. $F_{\nu_\Delta}(x) = 1 - (1+\alpha^{-1})^{-\alpha}(1-x)^\alpha$ if $-\alpha^{-1} \leq x \leq 1$. ∎

In each of these examples, $(\nu_\Delta)^* = \nu$. It is true that if $\nu \in \mathcal{P}^*$, then $\nu \prec_s (\nu_\Delta)^*$; but in general $\nu \neq (\nu_\Delta)^*$. Some examples in which $\nu \in \mathcal{P}^*$ and $\nu \neq (\nu_\Delta)^*$ are given in Example 2.9, and are a consequence of the following proposition. Let ν be any p.m. on $\mathbb{R}$. Define the set

$$\mathrm{Supp}_+(\nu) := \{x : \nu[y,\infty) < \nu[x,\infty) \text{ for all } y > x\},$$

and define the function $\Lambda(x;\nu)$ for $x \in \mathrm{Supp}_+(\nu)$ by

$$\Lambda(x) = \Lambda(x;\nu) = \inf_{y>x}(y\nu[y,\infty) - x\nu[x,\infty))/(\nu[y,\infty) - \nu(x,\infty)). \tag{2.18}$$

For a particular p.m. ν, this function $\Lambda(\cdot\,;\nu)$ was used in the proof of Proposition 2.1.

PROPOSITION 2.8: *Let ν be any discrete p.m. on $\mathbb{R}$ having finitely many atoms $y_0 < y_1 < \ldots < y_n$. Then a r.v. X which has minimal p.m. ν_Δ is given by*

$$(2.19)\qquad X = \begin{cases} \Lambda(x_0) & \text{with probability } \nu(-\infty, x_1), \\ \Lambda(x_{j-1}) & \text{with probability } \nu[x_{j-1}, x_j), \quad \text{for } j = 2, \ldots, k, \\ \Lambda(x_k) = y_n & \text{with probability } \nu[x_k, \infty) = \nu\{y_n\}, \end{cases}$$

where numbers $x_0 < x_1 < \cdots < x_k$ are chosen from $(y_i)_{i=0,\cdots,n}$ by the following:

(2.20) *Let $x_0 = y_0$.*
Having chosen $x_0, \cdots, x_j$, the next number x_{j+1}
is the maximal number $y_i > x_j$ for which
$\Lambda(x_j) = (y_i\nu[y_i, \infty) - x_j\nu[x_j, \infty))/(\nu[y_i, \infty) - \nu[x_j, \infty))$.
The procedure stops at $x_k = y_n$.

(If $\nu = \varepsilon_y$, then $\nu_\Delta = \varepsilon_y$ also.)

Proof: The degenerate case $\nu = \varepsilon_y$ immediately gives $\nu_\Delta = \varepsilon_y$, so we suppose that $n \geq 1$. Let λ denote the p.m. associated with r.v. X of (2.19). To obtain that λ is the minimal p.m. ν_Δ, we show

(2.21) (i) $\nu \prec_s \lambda^*$, and
(ii) if μ is any p.m. on $\mathbb{R}$ with $\displaystyle\int_0^\infty x\,d\mu < \infty$ and $\nu \prec_s \mu^*$,
then $\lambda^* \prec_s \mu^*$,

and then use Theorem 2.4. For this purpose, we denote $c_j := \Lambda(x_{j-1})$ for $j = 1, \ldots, k+1$, and note that from direct calculation (working through $F_\lambda, F_\lambda^{-1}$, and H_λ^{-1}), one obtains that maximal p.m. λ^* is given by

$$(2.22)\qquad \lambda^*[x, \infty) = \begin{cases} 1 & \text{if } x < x_0 = y_0, \\ \nu[x_j, \infty)(x_j - c_j)/(x - c_j) & \text{if } x_{j-1} \leq x < x_j, \quad \text{for } j = 1, \ldots, k, \\ \nu[x_k, \infty) = \nu\{y_n\} & \text{if } x = y_n = x_k, \quad \text{and} \\ 0 & \text{if } y_n < x. \end{cases}$$

We show that $\nu \prec_s \lambda^*$. First, observe from (2.22) that $\nu[x, \infty) = \lambda^*[x, \infty)$ for $x \leq y_0$ and for $y_n \leq x$. Now, for $x_{j-1} \leq x \leq x_j$, observe that (from (2.18))

$$\begin{aligned} c_j = \Lambda(x_{j-1}) &= (x_j\nu[x_j, \infty) - x_{j-1}\nu[x_{j-1}, \infty))/(\nu[x_j, \infty) - \nu[x_{j-1}, \infty)) \\ (2.23)\ &\leq (x\nu[x, \infty) - x_{j-1}\nu[x_{j-1}, \infty))/(\nu[x, \infty) - \nu[x_{j-1}, \infty)). \end{aligned}$$

Then use (2.22) and (2.23) to see that $\nu[x,\infty) \le \lambda^*[x,\infty)$ is equivalent to

$$\nu[x,\infty) \le \nu[x_{j-1},\infty)(x_{j-1}-c_j)/(x-c_j)$$

and to establish this last inequality.

Next, we show that (2.21)(ii) holds. Let μ be any p.m. on $\mathbb{R}$ with $\int_0^\infty x d\mu < \infty$ and $\nu \prec_s \mu^*$. We show that $\lambda^* \prec_s \mu^*$. First, observe that $\lambda^*[x,\infty) \le \mu^*[x,\infty) = 1$ for $x \le \int y d\mu$ (where $y_0 \le \int y d\mu$ since $\nu \prec_s \mu^*$), and $\lambda^*[x,\infty) = \nu[x,\infty) \le \mu^*[x,\infty)$ for $y_n \le x$. One shows that $\lambda^*[x,\infty) \le \mu^*[x,\infty)$ for $y_0 < x < y_n$ by proceeding from $x = \int y d\mu$ to $x = y_n$ as follows. For $\int y d\mu < x \le x_1$,

$$\text{(2.24)} \qquad \frac{x\lambda^*[x,\infty)-x_0}{\lambda^*[x,\infty)-1} = c_1 = \frac{x_1\lambda^*[x_1,\infty)-x_0}{\lambda^*[x_1,\infty)-1} = \frac{x_1\nu[x_1,\infty)-x_0}{\nu[x_1,\infty)-1}$$

$$\ge \frac{x_1\mu^*[x_1,\infty)-x_0}{\mu^*[x_1,\infty)-1} \ge \frac{x\mu^*[x,\infty)-x_0}{\mu^*[x,\infty)-1},$$

where the equalities follow from (2.22) and direct calculation; the first inequality follows from $\nu \prec_s \mu^*$ and the function $(x_1p - x_0)/(p-1)$ being decreasing in p on [0,1]; and the second inequality follows from the function

$$(x\mu^*[x,\infty)-x_0)/(\mu^*[x,\infty)-1)$$

being increasing in x (for example, use a change of variable $u = H_\mu(x)$ and $u_0 = H_\mu(x_0)$, the representation

$$(x\mu^*[x,\infty)-x_0)/(\mu^*[x,\infty)-1) = (u-u_0)^{-1}\int_{u_0}^{u} F^{-1}(t)dt$$

and calculus). It is immediate from (2.24) that $\lambda^*[x,\infty) \le \mu^*[x,\infty)$ for $x_0 < x \le x_1$.

Now suppose we know that $\lambda^*[x,\infty) \le \mu^*[x,\infty)$ for $x \le x_{j-1}$; we show that $\lambda^*[x,\infty) \le \mu^*[x,\infty)$ for $x_{j-1} \le x \le x_j$ in two cases. Let $p = \nu(-\infty, x_{j-1})$, and observe that $x_{j-1} \le H_\mu^{-1}(p)$ (from $\nu \prec_s \mu^*$). First, if $(\int y d\mu) \vee x_{j-1} \le x \le H_\mu^{-1}(p)$, then

$$\lambda^*[x,\infty) \le \lambda^*[x_{j-1},\infty) = \nu[x_{j-1},\infty) = 1-p = \mu^*[H_\mu^{-1}(p),\infty) \le \mu^*[x,\infty).$$

Second, suppose that $(\int y d\mu) \vee H_\mu^{-1}(p) \le x \le x_j$; we have that

$$(2.25)\quad \frac{x(\lambda^*[x,\infty)(1-p)^{-1}) - x_{j-1}}{(\lambda^*[x,\infty)(1-p)^{-1}) - 1} = c_j = \frac{x_j(\nu[x_j,\infty)(1-p)^{-1}) - x_{j-1}}{(\nu[x_j,\infty)(1-p)^{-1}) - 1}$$
$$\ge \frac{x_j(\mu^*[x_j,\infty)(1-p)^{-1}) - x_{j-1}}{(\mu^*[x_j,\infty)(1-p)^{-1}) - 1}$$
$$\ge \frac{x(\mu^*[x,\infty)(1-p)^{-1}) - x_{j-1}}{(\mu^*[x,\infty)(1-p)^{-1}) - 1},$$

where the reasoning is analogous to that used in verifying (2.24) with the exception of the last inequality. The last inequality of (2.25) uses that the function

$$g(x) := (x(\mu^*[x,\infty)(1-p)^{-1}) - x_{j-1})/((\mu^*[x,\infty)(1-p)^{-1}) - 1)$$
$$= (p-u)^{-1}\left(\int_u^1 F_\mu^{-1}(t)dt - x_{j-1}(1-p)\right)$$

is decreasing in x (here $u = H_\mu(x)$). This follows by observing that $dg/du \ge 0$ if and only if

$$(2.26)\qquad s(u) := \int_u^1 F_\mu^{-1}(t)dt + F_\mu^{-1}(u)(u-p) \ge x_{j-1}(1-p)$$

for $p \le u \le 1$. But $s(u)$ is increasing for $p \le u \le 1$, and so

$$s(u) \ge s(p) = (1-p)H_\mu^{-1}(p) \ge x_{j-1}(1-p),$$

and (2.26) is verified. From (2.25), it easily follows that $\lambda^*[x,\infty) \le \mu^*[x,\infty)$ for $H_\mu^{-1}(p) \le x \le x_j$. This completes the proof. ∎

Examples 2.9: The following examples show that for discrete p.m.'s ν with finitely many atoms, (a) the number of atoms of the associated minimal p.m.'s ν_Δ varies; and (b) in general, $\nu \ne (\nu_\Delta)^*$.

(i) For p.m. $\nu = (1/3)\varepsilon_0 + (1/3)\varepsilon_1 + (1/3)\varepsilon_2$, we have $\nu_\Delta = (1/3)\varepsilon_{-2} + (1/3)\varepsilon_0 + (1/3)\varepsilon_2$; and $(\nu_\Delta)^*[x,\infty) = 1$ if $x < 0, = 2(x+2)^{-1}$ if $0 \le x < 1, = 2(3x)^{-1}$ if $1 \le x < 2, = 1/3$ if $x = 2$, and $= 0$ if $2 < x$.

(ii) For p.m. $\nu = (1/3)\varepsilon_0 + (1/9)\varepsilon_1 + (5/9)\varepsilon_2$, we have $\nu_\Delta = (4/9)\varepsilon_{-5/2} + (5/9)\varepsilon_2$; and $(\nu_\Delta)^*[x,\infty) = 1$ if $x < 0, = 5(2x+5)^{-1}$ if $0 \le x < 2, = 5/9$ if $x = 2$, and $= 0$ if $2 < x$. ∎

Remark 2.10: We have given a characterization of the set $\mathcal{P}_0^*$ of maximal p.m.'s in Lemma 2.6 and of the set $\mathcal{P}^*$ of p.m.'s dominated above by maximal p.m.'s (in $\prec_s$ order) in Proposition 2.1. The analogous set of minimal p.m.'s and set of p.m.'s dominated below by minimal p.m.'s (in $\prec_c$ order) are equal to $\mathcal{E}_+$; that is,

$$(2.27) \qquad \begin{aligned} \mathcal{E}_+ &= \{\mu \in \mathcal{E}_+ : \text{ there is a p.m. } \nu \in \mathcal{P}^* \text{ with } \nu_\Delta \prec_c \mu\} \\ &= \{\mu \in \mathcal{E}_+ : \text{ there is a p.m. } \nu \in \mathcal{P}^* \text{ with } \nu_\Delta = \mu\}. \end{aligned}$$

This follows immediately from Lemma 2.5(a). ∎

3. Minimal p.m.'s, Maximal p.m.'s, and a Lattice Isomorphism

In this section, mappings $*$ and Δ are introduced to clarify the connection between minimal p.m.'s and maximal p.m.'s. Recall that $\mathcal{E}_+$ is the set of minimal p.m.'s (see Remark 2.10), and that $\mathcal{P}_0^*$ is the set of maximal p.m.'s (see (2.2)). Define the mappings $* : \mathcal{E}_+ \to \mathcal{P}_0^*$ and $\Delta : \mathcal{P}^* \to \mathcal{E}_+$ by $*(\mu) = \mu^*$ and $\Delta(\nu) = \nu_\Delta$. We show in Theorem 3.1 that $*$ is a lattice isomorphism between $(\mathcal{E}_+, \prec_c)$ and $(\mathcal{P}_0^*, \prec_s)$, where $\mathcal{E}_+$ has lattice operations $\vee_c$ and $\wedge_c$ of (1.4) and appropriate lattice operations on $\mathcal{P}_0^*$ are defined immediately before Theorem 3.1.

First, observe that mapping $*$ is a bijection between $\mathcal{E}_+$ and $\mathcal{P}_0^*$, with inverse mapping given by the restriction of Δ to $\mathcal{P}_0^*$. Indeed, use Lemma 1.8 and (2.7) to see that $*$ is one-to-one and maps $\mathcal{E}_+$ onto $\mathcal{P}_0^*$; and use Lemma 2.5(a) to obtain that the inverse of $*$ is the mapping Δ restricted to $\mathcal{P}_0^*$. In particular, for each maximal p.m. ν, there is only one $\mu \in \mathcal{E}_+$ for which $*(\mu) = \nu$. However, for each minimal p.m. $\mu \in \mathcal{E}_+$, there are many p.m.'s $\nu \in \mathcal{P}^*$ for which $\Delta(\nu) = \mu$. For example, let $\mu = (1-p)\varepsilon_a + p\varepsilon_b$, where $a < b$ and $0 < p < 1$ and denote

$$x_0 = \int x\,d\mu = (1-p)a + pb.$$

We claim that

$$\begin{aligned} &\{\nu \in \mathcal{P}(\mathbb{R}) : \nu_\Delta = \mu\} = \\ &\{\nu \in \mathcal{P}(\mathbb{R}) : x_0 = \inf\{x : \nu(-\infty, x] > 0\}, \nu[b,\infty) = \nu\{b\} = p \\ &\qquad \text{and } \nu[x,\infty) \le p(b-a)/(x-a) \text{ for } x_0 \le x \le b\}. \end{aligned}$$

To see that this holds, observe that for the given $\mu, \mu^*[x,\infty) = 1$ if $x < x_0, =$ $p(b-a)/(x-a)$ if $x_0 \le x \le b$, and $= 0$ if $b < x$; and use the defining properties of ν_Δ (see, e.g., Theorem 2.4). Thus, a restriction of Δ to $\mathcal{P}_0^*$ was necessary for description of the inverse of $*$.

A lattice structure on $\mathcal{P}_0^*$ is identified through the following operations. The operations $\wedge_*$ and $\vee_*$ are defined for pairs of p.m.'s σ_1 and σ_2 in $\mathcal{P}_0^*$ by

$$\sigma_1 \wedge_* \sigma_2 := ((\sigma_1 \wedge_s \sigma_2)_\Delta)^* \quad \text{and} \quad \sigma_1 \vee_* \sigma_2 := ((\sigma_1 \vee_s \sigma_2)_\Delta)^*.$$

Here are some easily verified observations related to this definition: (i) $\sigma_1 \wedge_s \sigma_2 \prec_s \sigma_1 \wedge_* \sigma_2$ and $\sigma_1 \vee_s \sigma_2 \prec_s \sigma_1 \vee_* \sigma_2$; (ii) $\sigma_1 \wedge_* \sigma_2$ and $\sigma_1 \vee_* \sigma_2$ are in $\mathcal{P}_0^*$; and (iii) if $\sigma_1 = \mu_1^*$ and $\sigma_2 = \mu_2^*$ for μ_1 and μ_2 in $\mathcal{E}_+$, then

$$\sigma_1 \wedge_* \sigma_2 = \mu_1^* \wedge_* \mu_2^* = (\mu_1 \wedge_c \mu_2)^* \quad \text{and} \; \sigma_1 \vee_* \sigma_2 = \mu_1^* \vee_* \mu_2^* = (\mu_1 \vee_c \mu_2)^*.$$

It is easy to see that $(\mathcal{P}_0^*, \prec_s)$ is a lattice under the operations $\wedge_*$ and $\vee_*$.

THEOREM 3.1: *The map* $* : (\mathcal{E}_+, \prec_c, \wedge_c, \vee_c) \to (\mathcal{P}_0^*, \prec_s, \wedge_*, \vee^*)$ *is a lattice isomorphism. The inverse of* $*$ *is the map restricted to* $\mathcal{P}_0^*$.

Proof: It remains only to prove the lattice preserving property, that is,

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

But this is a straightforward consequence of Lemma 1.8 and the definitions of the $*$ and Δ operations. ∎

As a consequence of Theorem 3.1, we have the following characterization of operations $\vee_c$ and $\wedge_c$ in terms of minimal p.m.'s and maximal p.m.'s: for μ_1 and μ_2 in $\mathcal{E}_+$,

$$\mu_1 \wedge_c \mu_2 = \Delta(\mu_1^* \wedge_s \mu_2^*) \quad \text{and} \quad \mu_1 \vee_c \mu_2 = \Delta(\mu_1^* \vee_s \mu_2^*). \tag{3.1}$$

The following example shows that $\mathcal{P}_0^*$ is not a lattice under the operations $\vee_s$ and $\wedge_s$ of Section 1.

Example 3.2: We exhibit p.m.'s μ_1 and μ_2 in $\mathcal{E}_+$ (so μ_1^* and μ_2^* are in $\mathcal{P}_0^*$) for which

$$\mu_1^* \vee_s \mu_2^* \notin \mathcal{P}_0^* \quad \text{and} \quad \mu_1^* \vee_s \mu_2^* \not\approx_s ((\mu_1^* \vee_s \mu_2^*)_\Delta)^* = \mu_1^* \vee_* \mu_2^*.$$

Let w be a uniform distributed r.v. on [0,1], and define

$$\mu_1 \overset{\mathcal{D}}{=} F_1^{-1}(w) = \begin{cases} 2 & \text{if } w \in (\frac{1}{3}, 1] \\ -4 & \text{if } w \in [0, \frac{1}{3}] \end{cases}$$

and

$$\mu_2 \overset{\mathcal{D}}{=} F_2^{-1}(w) = \begin{cases} 4 & \text{if } w \in (\frac{2}{3}, 1], \\ -2 & \text{if } w \in [0, \frac{2}{3}]. \end{cases}$$

Then

$$\mu_1^* \overset{\mathcal{D}}{=} H_1^{-1}(w) = \begin{cases} 2 & \text{if } w \in (\frac{1}{3}, 1] \\ 4w/(1-w) & \text{if } w \in [0, \frac{1}{3}] \end{cases}$$

and

$$\mu_2^* \overset{\mathcal{D}}{=} H_2^{-1}(w) = \begin{cases} 4 & \text{if } w \in (\frac{2}{3}, 1] \\ 2w/(1-w) & \text{if } w \in [0, \frac{2}{3}] \end{cases}$$

and

$$\mu_1^* \vee_s \mu_2^* \overset{\mathcal{D}}{=} (H_1^{-1} \vee H_2^{-1})(w) = \begin{cases} 4 & \text{if } w \in (\frac{2}{3}, 1] \\ 2w/(1-w) & \text{if } w \in (\frac{1}{2}, \frac{2}{3}] \\ 2 & \text{if } w \in (\frac{1}{3}, \frac{1}{2}] \\ 4w/(1-w) & \text{if } w \in (0, \frac{1}{3}]. \end{cases}$$

Next, denote p.m. ρ on $\mathbb{R}$ by

$$\rho \overset{\mathcal{D}}{=} M = \begin{cases} 4 & \text{if } w \in (\frac{2}{3}, 1] \\ 2 & \text{if } w \in (\frac{1}{3}, \frac{2}{3}] \\ 0 & \text{if } w \in [0, \frac{1}{3}]. \end{cases}$$

Then

$$\rho_\Delta \overset{\mathcal{D}}{=} F^{-1}(w) = \begin{cases} 4 & \text{if } w \in (\frac{2}{3}, 1] \\ 0 & \text{if } w \in (\frac{1}{3}, \frac{2}{3}] \\ -4 & \text{if } w \in [0, \frac{1}{3}] \end{cases}$$

and

$$(\rho_\Delta)^* \overset{\mathcal{D}}{=} H^{-1}(w) = \begin{cases} 4 & \text{if } w \in (\frac{2}{3}, 1] \\ (4/3)/(1-w) & \text{if } w \in (\frac{1}{3}, \frac{2}{3}] \\ 4w/(1-w) & \text{if } w \in [0, \frac{1}{3}]. \end{cases}$$

It follows that

$$\rho \precnsim_s \mu_1^* \vee_s \mu_2^* \precnsim_s (\rho_\Delta)^* = ((\mu_1^* \vee_s \mu_2^*)_\Delta)^* = \mu_1^* \vee_* \mu_2^*. \tag{3.2}$$

In particular, (3.2) indicates that $\mu_1^* \vee_s \mu_2^* \notin \mathcal{P}_0^*$. (From Lemma 1.8, we have also that $\rho_\Delta = (\mu_1^* \vee_s \mu_2^*)_\Delta$.)

4. Some Properties of Minimal Probability Measures

In this section, we prove several important properties relating probability measures and their minimal probability measures. These properties will be used in Section 5.

PROPOSITION 4.1: *For each $\nu \in \mathcal{P}^*$, the minimal p.m. ν_Δ satisfies $\nu_\Delta \prec_s \nu$.*

Proof: Denote the distribution functions of ν_Δ and ν by F_Δ and F respectively. Suppose there is $u_0 \in (0,1)$ with $F^{-1}(u_0) < F_\Delta^{-1}(u_0)$. We obtain a contradiction to the minimality of ν_Δ (i.e., $\nu_\Delta = \text{g.l.b.}_{\prec_c} \mathcal{E}_+(\nu)$). Define

$$u_2 := \inf\{u > u_0 : F_\Delta^{-1}(u) \le F^{-1}(u)\}$$

if this set $\neq \emptyset$, and $= 1$ otherwise. By the left continuity of the left-inverses, we may assume $u_0 < u_2$; we have $F^{-1}(u) < F_\Delta^{-1}(u)$ for $u \in (u_0, u_2)$. Next define

$$u_1 := \sup\left\{ t < u_2 : \int_t^{u_2} F_\Delta^{-1}(u)du < \int_t^{u_2} F^{-1}(u)du \right\}$$

if this set $= \emptyset$, and $= 1$ otherwise. Then $u_1 < u_0 < u_2$;

$$\int_t^{u_2} F^{-1}(u)du \le \int_t^{u_2} F_\Delta^{-1}(u)du \quad \text{for } t \in (u_1, u_2);$$

and if $u_1 > 0$, then

$$\int_{u_1}^{u_2} F_\Delta^{-1}(u)du = \int_{u_1}^{u_2} F^{-1}(u)du \quad \text{and} \quad F_\Delta^{-1}(u_1+) \le F^{-1}(u_1+)$$

(use for all $\varepsilon > 0$ sufficiently small,

$$\int_{u_1+\varepsilon}^{u_2} (F_\Delta^{-1}(u) - F^{-1}(u))du \ge 0 \quad \text{and} \quad \int_{u_1-\varepsilon}^{u_2} (F_\Delta^{-1}(u) - F^{-1}(u))du \le 0).$$

Define $\tilde{F}^{-1}$ on [0,1] by

$$\tilde{F}^{-1}(u) = \begin{cases} F_\Delta^{-1}(u) & \text{if } 0 \le u \le u_1 \\ F^{-1}(u) & \text{if } u_1 < u \le u_2 \\ F_\Delta^{-1}(u) & \text{if } u_2 < u \le 1 \end{cases}$$

(if $u_2 = 1$, define $\tilde{F}^{-1}(u) = F^{-1}(u)$ for $u_1 < u \le 1$; and if $u_1 = 0$, define $\tilde{F}^{-1}(u) = F^{-1}(u)$ for $0 \le u \le u_2$). Let $\tilde{\nu}$ be the associated p.m. and $\tilde{F}$ be its distribution function; then $\tilde{F}^{-1}$ is the associated left inverse of $\tilde{F}$. We claim that

$$\tilde{\nu} \in \mathcal{E}_+(\nu), \quad \tilde{\nu} \prec_c \nu_\Delta, \quad \text{and} \quad \tilde{\nu} \neq \nu_\Delta. \tag{4.1}$$

This will give the desired contradiction to the minimality of ν_Δ. To see that $\tilde{\nu} \in \mathcal{E}_+(\nu)$, first observe that $\tilde{\nu} \in \mathcal{E}_+$ since $\nu_\Delta \in \mathcal{E}_+$, and that $\nu \prec_s (\tilde{\nu})^*$ will follow if we show $F^{-1} \le \tilde{H}^{-1}$. But this is immediate from $F^{-1} \le H_\Delta^{-1}$ $\quad (\nu \prec_s (\nu_\Delta)^*)$ and

$$\tilde{H}^{-1}(t) = \begin{cases} H_\Delta^{-1}(t) & \text{if } 0 \le t \le u_1, \\ (1-t)^{-1}\{\int_t^{u_2} F^{-1}(u)du + (1-u_2)H_\Delta^{-1}(u_2)\} & \text{if } u_1 \le t \le u_2, \\ H_\Delta^{-1}(t) & \text{if } u_2 \le t \le 1. \end{cases}$$

Next, $\tilde{\nu} \prec_c \nu_\Delta$ follows from Lemma 1.8 once it is shown that $\tilde{H}^{-1} \le H_\Delta^{-1}$. But this is immediate for $t \le u_1$ and for $t \ge u_2$, and for $u_1 \le t \le u_2$ it follows from

$$\begin{aligned} \tilde{H}^{-1}(t) &= (1-t)^{-1}\left\{ \int_t^{u_2} F^{-1}(u)du + (1-u_2)H_\Delta^{-1}(u_2) \right\} \\ &\le (1-t)^{-1}\left\{ \int_t^{u_2} F_\Delta^{-1}(u)du + (1-u_2)H_\Delta^{-1}(u_2) \right\} = H_\Delta^{-1}(t). \end{aligned}$$

Finally $\tilde{\nu} \neq \nu_\Delta$, since for $u \in (u_0, u_2)$, $\quad \tilde{F}^{-1}(u) = F^{-1}(u) < F_\Delta^{-1}(u)$. This gives that (4.1) holds. ∎

The following lemma is useful for establishing properties satisfied by minimal p.m.'s, as in the proof of Theorem 4.5. Recall from Section 2 that '$\Rightarrow$' denotes weak convergence.

LEMMA 4.2: (a) *Let $\nu \in \mathcal{P}^*$ and let $\{\nu_n\}_{n\in\mathbb{N}}$ be a sequence of p.m.'s on $\mathbb{R}$ which increases to ν in the $\prec_s$ order. Then the sequence $\{((\nu_n)_\Delta)^*\}$ is increasing in the $\prec_s$ order; $((\nu_n)_\Delta)^* \Rightarrow (\nu_\Delta)^*; (\nu_\Delta)^* = \vee_s\{((\nu_n)_\Delta)^* : n \in \mathbb{N}\}$; and $\nu_\Delta = \vee_c\{(\nu_n)_\Delta : n \in \mathbb{N}\}$.*

(b) *Let $\nu_1 \in \mathcal{P}^*$ and $\{\nu_n\}_{n\in\mathbb{N}}$ be a sequence of p.m.'s on $\mathbb{R}$ which decreases to p.m. ν in the $\prec_s$ order. Then the sequence $\{((\nu_n)_\Delta)^*\}$ is decreasing in the $\prec_s$ order; $((\nu_n)_\Delta)^* \Rightarrow (\nu_\Delta)^*; (\nu_\Delta)^* = \wedge_s\{((\nu_n)_\Delta)^* : n \in \mathbb{N}\}$; and $\nu_\Delta = \wedge_c\{(\nu_n)_\Delta : n \in \mathbb{N}\}$.*

Proof: We prove part (a); the proof of part (b) is analogous. Let $\mu_0 \in \mathcal{E}_+$ satisfying $\nu \prec_s \mu_0^*$. We are given

(4.2)

$$\nu_n \prec_s \nu_{n+1} \prec_s \nu \quad \text{for all } n \in \mathbb{N}, \quad \text{and} \quad \lim_n \nu_n[x,\infty) = \nu[x,\infty) \quad \text{for all } x \in \mathbb{R}.$$

From Theorem 2.4, one obtains the sequence $\{(\nu_n)_\Delta\}$ of p.m.'s in $\mathcal{E}_+$ and uses their properties to show that $((\nu_1)_\Delta)^* \prec_s ((\nu_n)_\Delta)^* \prec_s \mu_0^*$ for all $n \in \mathbb{N}$ and that the sequence $\{((\nu_n)_\Delta)^*\}$ is increasing in the $\prec_s$ order. Then apply Lemma 2.3(b) to obtain a unique p.m. $\mu \in \mathcal{E}_+$ satisfying $((\nu_n)_\Delta)^* \Rightarrow \mu^*$ and $\mu^* = \vee_s\{((\nu_n)_\Delta)^* : n \in \mathbb{N}\}$. Now apply Theorem 2.4 once again to obtain the p.m. ν_Δ in $\mathcal{E}_+$. Since $\nu_n \prec_s \nu \prec_s (\nu_\Delta)^*$, it follows that $((\nu_n)_\Delta)^* \prec_s (\nu_\Delta)^*$, for all $n \in \mathbb{N}$; and hence $\mu^* \prec_s (\nu_\Delta)^*$. Also, $\nu_n \prec_s ((\nu_n)_\Delta)^* \prec_s \mu^*$ for all $n \in \mathbb{N}$, and $\nu \prec_s \mu^*$ follows from (4.2); and hence $(\nu_\Delta)^* \prec_s \mu^*$. Thus $\mu = \nu_\Delta$. The remainder of the conclusion follows from Lemma 1.8. ∎

In general, additional care is needed in combining convergence of p.m.'s $\{\mu_n^*\}$ in $\prec_s$-order with convergence of means $\{\int x d\mu_n\}_{n\geq 1}$. For this purpose, first recall from ([3]: page 92 and [4]: page 89) that $\{\mu_n\}_{n\in\mathbb{N}}$ is uniformly integrable from below if

$$\lim_{A\to\infty} \sup_n \int_{x\leq -A} -x d\mu_n = 0.$$

We state convergence results for means in the following lemma and then apply them in Theorem 4.5.

LEMMA 4.3: *Let $\{\mu_n\}_{n\in\mathbb{N}}$ be any sequence of p.m.'s in $\mathcal{E}_+$ satisfying $\nu \prec_s \mu_n^* \prec_s \mu_0^*$ for all $n \in \mathbb{N}$, for some $\nu \in \mathcal{P}(\mathbb{R})$ and $\mu_0 \in \mathcal{E}_+$. Then*

- (a) *if $\{\mu_n^*\}$ is decreasing in the $\prec_s$ order and μ is the p.m. of Lemma 2.3(a), then $\int x d\mu_n \downarrow \int x d\mu$ as $n \to \infty$, in $[-\infty,\infty)$; and*
- (b) *if $\{\mu_n^*\}$ is increasing in the $\prec_s$ order and μ is the p.m. of Lemma 2.3(b), then $\lim_n \uparrow \int x d\mu_n \leq \int x d\mu$. If, in addition, $\{\mu_n\}$ is uniformly integrable from below, then $\int x d\mu_n \uparrow \int x d\mu$ as $n \to \infty$.*

Proof: Under the assumptions of part (a), first observe that $\{\int x d\mu_n\}$ is a nonincreasing sequence in $[-\infty, \infty)$, since

$$(4.3)\qquad \int x d\mu_n = \int_0^1 F_{\mu_n}^{-1}(u)du = \lim_{t\downarrow 0} \int_t^1 F_{\mu_n}^{-1}(u)du = \lim_{t\downarrow 0}(1-t)H_{\mu_n}^{-1}(t)$$

$$\geq \lim_{t\downarrow 0}(1-t)H_{\mu_{n+1}}^{-1}(t) = \int_0^1 F_{\mu_{n+1}}^{-1}(u)du = \int x d\mu_{n+1}.$$

We may interchange limits as follows to obtain

$$\lim_n \int x d\mu_n = \lim_n \int_0^1 F_{\mu_n}^{-1}(u)du = \lim_n \lim_{t\downarrow 0} H_{\mu_n}^{-1}(t)$$

$$= \lim_{t\downarrow 0} \lim_n H_{\mu_n}^{-1}(t) = \lim_{t\downarrow 0} H_{\mu}^{-1}(t) = \int_0^1 F_{\mu}^{-1}(u)du = \int x d\mu.$$

Under the assumptions of part (b), one has that $\{\int x d\mu_n\}$ is a nondecreasing sequence in $[-\infty, \infty)$ (argue analgous to (4.3)). Also, for every $t \in (0,1)$, we have

$$\int x d\mu_n = \int_0^1 F_{\mu_n}^{-1}(u)du = H_{\mu_n}^{-1}(0) \leq H_{\mu_n}^{-1}(t) \leq H_{\mu}^{-1}(t),$$

and thus

$$(4.4)\qquad \lim_n \int x d\mu_n \leq \lim_{t\downarrow 0} H_{\mu}^{-1}(t) = H_{\mu}^{-1}(0) = \int_0^1 F_{\mu}^{-1}(u)du = \int x d\mu.$$

If, in addition, $\{\mu_n\}$ is uniformly integrable from below, and $\{\mu_{n_i}\}$ is a subsequence of $\{\mu_n\}$ for which $\mu_{n_i} \Rightarrow \mu$ (as in the proof of Lemma 2.2), then Fatou's Lemma can be applied (e.g., in the form given in [3]: page 94) to give

(4.5)

$$\int x d\mu = \int_0^1 F^{-1}(u)du = \int_0^1 \lim_{n_i} F_{n_i}^{-1}(u)du \leq \liminf_n \int_0^1 F_{n_i}^{-1}(u)du = \lim_n \int x d\mu$$

and equality follows from (4.4) and (4.5). ∎

The following example illustrates that one must add some additional hypothesis to those of part (b) of Lemma 4.3, such as uniform integrability of $\{\mu_n\}$ from below, in order to obtain convergence of $\{\int x d\mu_n\}$ to $\int x d\mu$.

Example 4.4: For $n \in \mathbb{N}$, let $\mu_n = n^{-1}\varepsilon_{-n} + (1-n^{-1})\varepsilon_0$; and let $\mu_0 = \varepsilon_0$ and $\nu = \varepsilon_{-1}$. Then for $n \in \mathbb{N}$, μ_n and μ_n^* have respective left-inverse functions F_n^{-1} and H_n^{-1} given by $F_n^{-1}(u) = -n$ if $0 \le u \le n^{-1}$, and $= 0$ if $n^{-1} < u \le 1$; and $H_n^{-1}(t) = (-n)(n^{-1}-t)/(1-t)$ if $0 \le t \le n^{-1}$, and $= 0$ if $n^{-1} < t \le 1$. The assumptions under part (b) of Lemma 4.3 are satisfied, and $\mu = \varepsilon_0$. However, $\lim_n \int x d\mu_n = -1 < 0 = \int x d\mu$.

THEOREM 4.5: *For each $\nu \in \mathcal{P}^*$,*

$$\min_{\mu \in \mathcal{E}_+(\nu)} \int x d\mu = \int x d\nu_\Delta = \inf\{z : \nu(-\infty, z] > 0\}, \tag{4.6}$$

and this number is in $[-\infty, \infty)$.

Proof: Let $\nu \in \mathcal{P}^*$, so that there is a p.m. $\mu_0 \in \mathcal{E}_+$ with $\nu \prec_s \mu_0^*$. The first equality in (4.6) is clear, e.g., from (2.14)(i). We divide the proof of the second equality into three steps.

STEP 1: Assume ν is a discrete p.m. having finitely many atoms $x_0 < \cdots < x_n$. Use Proposition 2.8, and, in particular, (2.19) and (2.23), to obtain that $\int x d\nu_\Delta = x_0$ from direct computation.

STEP 2: Let $x_0 = \inf\{z : \nu(-\infty, z] > 0\}$ and assume x_0 is finite with $\nu\{x_0\} > 0$. Let D be some dense subset of $[x_0, \infty)$ containing the atoms of ν. Choose a scheme $x_{n,i} \in D, i = 0, \ldots, n, n \in \mathbb{N}$, of real numbers with $x_{n,0} = x_0 < x_{n,1} < \ldots < x_{n,n}$ and $D_n := \{x_{n,i} : i = 0, \ldots, n\}$ increasing under containment to D as $n \to \infty$. Define discrete p.m.'s $\nu_n, n \in \mathbb{N}$, by $\nu_n[x, \infty) = 1$ if $x \le x_{n,0} = x_0, =$ $\nu[x_{n,i}, \infty)$ if $x_{n,i-1} < x \le x_{n,i}$, for $i = 1, \ldots, n$, and $= 0$ if $x_{n,n} < x$. The ν_n's are well-defined and satisfy by construction that $\{\nu_n\}$ increases to ν in the $\prec_s$ order. Thus $\nu_n \in \mathcal{P}^*$ for each $n \in \mathbb{N}$. From Lemma 4.2, we have $\{((\nu_n)_\Delta)^*\}$ increases in the $\prec_s$ order and $(\nu_\Delta)^* = \vee_s\{((\nu_n)_\Delta)^* : n \in \mathbb{N}\}$. Also, $\int x d(\nu_n)_\Delta = x_0$ for all $n \in \mathbb{N}$ from Step 1; and $\{(\nu_n)_\Delta\}$ is uniformly integrable from below. In fact, Proposition 2.8 gives that the support of $(\nu_n)_\Delta$ is contained in $[\alpha, \infty)$ for all $n \in \mathbb{N}$ where

$$-\infty < \alpha := \inf_{y > x_0} (y\nu[y, \infty) - x_0)/(\nu[y, \infty) - 1)$$

$$\leq \inf_{y>x_0}(y\nu_n[y,\infty)-x_0)/(\nu_n[y,\infty)-1).$$

From Lemma 4.3(b), it follows that $\lim_n \int xd(\nu_n)_\Delta = \int xd\nu_\Delta$, and so $\int xd\nu_\Delta = x_0$.

STEP 3: For any $\nu \in \mathcal{P}^*$, let $x_0 := \inf\{z : \nu(-\infty, z] > 0\}$ and assume $\nu\{x_0\} = 0$. In this case, x_0 might be $-\infty$. Let $\{\delta_n\}_{n\in\mathbb{N}}$ be a sequence of numbers decreasing to x_0, and define p.m.'s $\nu^n, n \in \mathbb{N}$, by

$$\nu^n(\cdot) = \nu(\cdot \cap [\delta_n, \infty)) + \nu(-\infty, \delta_n)\varepsilon_{\delta_n}(\cdot).$$

Then $\nu^n \in \mathcal{P}^*$ and $\{\nu^n\}$ decreases to ν in the $\prec_s$ order. From Lemma 4.2, we have $\{((\nu^n)_\Delta)^*\}$ decreases in the $\prec_s$ order and

$$(\nu_\Delta)^* = \wedge_s\{((\nu^n)_\Delta)^* : n \in \mathbb{N}\}.$$

Also, $\int xd(\nu^n)_\Delta = \delta_n$ for each $n \in \mathbb{N}$ from Step 2. From Lemma 4.3(a), it follows that $\lim_n \int xd(\nu^n)_\Delta = \int xd\nu_\Delta$, and so $\int xd\nu_\Delta = x_0 = \inf\{z : \nu(-\infty, x] > 0\}$. ∎

5. Martingale Representations and Inequalities

In this section, we give several characterizations of collections of martingales and martingale inequalities which use minimal p.m.'s. In particular, sharp, expectation-based martingale inequalities are given in Theorem 5.3. In Remark 5.4 the inequalities in Theorem 5.3 are given a natural interpretation as a new class of 'prophet inequalities' in optimal stopping theory.

In this section we use the following definition of an (integrable) martingale. $X = (X_t)_{0\leq t\leq 1}$ is a **martingale** if there is some probability space $(\Omega, \mathcal{F}, P)$ and a filtration $\{\mathcal{F}_t\}_{0\leq t\leq 1}$ on $(\Omega, \mathcal{F}, P)$ under which (i) $(X_t)_{0\leq t\leq 1}$ is $\{\mathcal{F}_t\}$-adapted; (ii) $E|X_t| < \infty$ for every $0 \leq t \leq 1$; (iii) $E(X_t|\mathcal{F}_s) = X_s$ a.e. [P] for every $0 \leq s \leq t \leq 1$; and (iv) the paths $t \mapsto X_t$ are right continuous and have left-hand limits in $\mathbb{R}$ for $0 \leq t \leq 1$ (RCCL).

We recall the following results for martingales $X = (X_t)_{0\leq t\leq 1}$ which relate the terminal r.v. X_1 and the supremum r.v. $M = M(X) = \sup_{0\leq t\leq 1} X_t$ ([8]: Theorems 2.1 and 3.1).

Let μ be any p.m. on $\mathbb{R}$ with $\int xd\mu(x)$ finite. Then

$$\{\nu \in \mathcal{P}(\mathbb{R}) : \text{there is a martingale } (X_t)_{0\leq t\leq 1} \text{ satisfying } M \overset{\mathcal{D}}{=} \nu \text{ and } X_1 \overset{\mathcal{D}}{=} \mu\}$$
$$= \{\nu \in \mathcal{P}(\mathbb{R}) : \mu \prec_s \nu \prec_s \mu^*\}. \tag{5.1}$$

Let μ be any p.m. on $\mathbb{R}$ with $\int x d\mu(x) = x_0 \in \mathbb{R}$. Then

$$\text{(5.2)} \quad \begin{aligned} &\{\nu \in \mathcal{P}(\mathbb{R}) : \text{ there is a martingale } (X_t)_{0 \le t \le 1} \text{ satisfying} \\ &X_0 \equiv x_0, M \overset{\mathcal{D}}{=} \nu \text{ and} X_1 \overset{\mathcal{D}}{=} \mu\} = \{\nu \in \mathcal{P}([x_0, \infty)) : \mu \prec_s \nu \prec_s \mu^*\}. \end{aligned}$$

We establish 'converses' of these results with use of maximal p.m.'s and minimal p.m.'s.

THEOREM 5.1: *Let ν be any p.m. in $\mathcal{P}^*$. Then*

$$\begin{aligned} &\{\mu \in \mathcal{P}(\mathbb{R}) : \textit{ there is a martingale } (X_t)_{0 \le t \le 1} \textit{ satisfying } M \overset{\mathcal{D}}{=} \nu \textit{ and } X_1 \overset{\mathcal{D}}{=} \mu\} \\ \text{(5.3)} \qquad &= \{\mu \in \mathcal{E} : \mu \prec_s \nu \prec_s \mu^*\} = \{\mu \in \mathcal{E} : \nu_\Delta \prec_c \mu \prec_s \nu\}. \end{aligned}$$

If, also, $x_0 := \inf\{z : \nu(-\infty, z] > 0\}$ is finite, then

$$\text{(5.4)} \quad \begin{aligned} \{\mu \in \mathcal{P}(\mathbb{R}) : &\text{there is a martingale } (X_t)_{0 \le t \le 1} \quad \text{satisfying} \\ &X_0 \equiv x_0, M \overset{\mathcal{D}}{=} \nu \text{ and } X_1 \overset{\mathcal{D}}{=} \mu\} \\ &= \left\{\mu \in \mathcal{P}(\mathbb{R}) : \mu \prec_s \nu \prec_s \mu^* \text{ and; } \int x d\mu = x_0\right\} \\ &= \{\mu \in \mathcal{E} : \nu_\Delta \prec_k \mu \prec_s \nu\}. \end{aligned}$$

Proof: We prove (5.4); the proof of (5.3) is analogous. Let $\nu \in \mathcal{P}^*$ with finite $x_0 = \inf\{z : \nu(-\infty, z] > 0\}$. First, assume that μ is a p.m. on $\mathbb{R}$ with $\int x d\mu = x_0$, and there is a martingale $X = (X_t)_{0 \le t \le 1}$ satisfying $X_0 \equiv x_0, M(X) \overset{\mathcal{D}}{=} \nu$, and $X_1 \overset{\mathcal{D}}{=} \mu$. It follows from (5.2) that

$$\text{(5.5)} \qquad \mu \prec_s \nu \prec_s \mu^* \quad \text{and} \quad \int x d\mu = x_0.$$

From Theorems 2.4 and 4.5, we have that $(\nu_\Delta)^* \prec_s \mu^*$ and $\int x d\nu_\Delta = x_0 = \int x d\mu$. It follows from (1.5) and (5.5) that

$$\text{(5.6)} \qquad \nu_\Delta \prec_k \mu \prec_s \nu \qquad \text{and} \qquad \int x d\mu = x_0.$$

Thus, we have established containment in one direction for (5.4).

Now, if μ is a p.m. on $\mathbb{R}$ satisfying (5.6), then, from (1.5) and Theorems 2.4 and 4.5, we have $\nu \prec_s (\nu_\Delta)^* \prec_s \mu^*$, and thus μ satisfies (5.5). But since μ is a p.m. on $\mathbb{R}$ satisfying (5.5), one then has, from (5.2), that there is a martingale $X = (X_t)_{0 \le t \le 1}$ satisfying $X_0 \equiv x_0, M(X) \overset{\mathcal{D}}{=} \nu$ and $X_1 \overset{\mathcal{D}}{=} \mu$. This proves containment in the other direction for (5.4). ∎

Observe that ν_Δ might not be in the set of p.m.'s in (5.3). Indeed, if $\nu \in \mathcal{P}^*$ with $x_0 := \inf\{z : \nu(-\infty, z] > 0\}$, and $x_0 = -\infty$, then $\int x d\nu_\Delta = -\infty$ (from Theorem 4.5), so $\nu_\Delta \notin \mathcal{E}$ and ν_Δ is not in the set in (5.3). On the other hand, x_0 is assumed to be finite in the second conclusion of Theorem 5.1, and ν_Δ is in the set of p.m.'s in (5.4). This leads to the following sharp results for (integrable) martingales in the setting of (5.4).

Using properties and characterizations of the Hardy and Littlewood maximal p.m., Dubins and Gilat [6] have shown that for any p.m. μ on $\mathbb{R}$ with $\int x d\mu = x_0$, the maximal p.m. μ^* satisfies $\mu^* \in \mathcal{M}(\mu; x_0)$ and $\mu^* =$ l.u.b. $\mathcal{M}(\mu; x_0)$ in the $\prec_s$ order, where

$$\mathcal{M}(\mu; x_0) := \{\nu \in \mathcal{P}(\mathbb{R}) : \text{there is a martingale } (X_t)_{0\le t\le 1} \text{ with } X_0 \equiv x_0, M \overset{\mathcal{D}}{=} \nu, \text{ and } X_1 \overset{\mathcal{D}}{=} \mu\}.$$

We have an analogous characterization result for minimal p.m.'s as a consequence of Theorem 5.1.

COROLLARY 5.2: *Let $\nu \in \mathcal{P}^*$ with finite $x_0 = \inf\{z : \nu(-\infty, z] > 0\}$, and let ν_Δ be the minimal p.m. associated with ν. Then $\nu_\Delta \in \mathcal{N}(\nu; x_0)$ and $\nu_\Delta =$ g.l.b. $\mathcal{N}(\nu; x_0)$ in the $\prec_k$ order, where*

$$\mathcal{N}(\nu; x_0) := \{\mu \in \mathcal{P}(\mathbb{R}) : \text{there is a martingale } (X_t)_{0\le t\le 1} \text{ with } X_0 \equiv x_0, M \overset{\mathcal{D}}{=} \nu, \text{ and } X_1 \overset{\mathcal{D}}{=} \mu\}.$$

Proof: We have that $\nu_\Delta \in \mathcal{N}(\nu; x_0)$ from Theorems 2.4 and 4.5, Proposition 4.1 and (5.4). To see that $\nu_\Delta =$ g.l.b. $\mathcal{N}(\nu; x_0)$ in the $\prec_k$ order, first observe that $\nu_\Delta \prec_k \mu$ for every $\mu \in \mathcal{N}(\nu; x_0)$ from (5.4); and if ρ is any p.m. in $\mathcal{E}$ satisfying $\rho \prec_k \mu$ for every $\mu \in \mathcal{N}(\nu; x_0)$, then $\rho \prec_k \nu_\Delta$, since $\nu_\Delta \in \mathcal{N}(\nu; x_0)$. ∎

The following martingale inequality form of Corollary 5.2 is a consequence of the definition of the $\prec_k$ order.

THEOREM 5.3: *Let $\nu \in \mathcal{P}^*$ with finite $x_0 = \inf\{z : \nu(-\infty, z] > 0\}$, and let ν_Δ be the minimal p.m. associated with ν. Then for every martingale $(X_t)_{0\le t\le 1}$ satisfying $X_0 = x_0, \sup_{0\le t\le 1} X_t \overset{\mathcal{D}}{=} \nu$, and with terminal r.v. X_1, it follows that*

$$\int \phi d\nu_\Delta \le E(\phi(X_1)) \quad \text{for each convex function } \phi \text{ on } \mathbb{R}, \tag{5.7}$$

provided both integrals exist. Within this collection of martingales, there is one with $X_1 \overset{\mathcal{D}}{=} \nu_\Delta$, and thus the inequalities in (5.7) are attained simultaneously.

Remark 5.4: The result in Theorem 5.3 has the following 'gambler vs. prophet' interpretation. Consider the collection of martingale games $X = (X_t)_{0\le t\le 1}$ starting at $X_0 \equiv x_0$. Among these games, consider the subcollection of martingale games for which $M = M(X) = \sup_{0\le t\le 1} X_t$ has distribution ν in $\mathcal{P}^*$. For each martingale game $(X_t)_{0\le t\le 1}$ within this subcollection, and any nondecreasing convex utility function ϕ, players of this game attempt to maximize their reward $E(\phi(X_\tau))$ through a well chosen stopping time τ. A 'prophet' playing this game, with anticipatory stop rules allowed, obtains reward $E(\phi(M)) = \int \phi d\nu$; and a 'gambler' playing this game, with only nonanticipatory stop rules allowed, obtains reward $E(\phi(X_1))$. *Theorem 5.3 states that within this subcollection of games, the gambler does worst under the martingale game $(X_t)_{0\le t\le 1}$ having distribution ν_Δ for the terminal r.v. X_1, and in this game the gambler obtains reward $E(\phi(X_1)) = \int \phi d\nu_\Delta$.* (Compare to Section 4 of [8]; see also [2].)

If $\nu \in \mathcal{P}^*$ with $x_0 = \inf\{z : \nu(-\infty, z] > 0\}$ finite, then ν_Δ is in the set of p.m.'s of (5.4), and so there exists a martingale $(Y_t)_{0\le t\le 1}$ satisfying $Y_0 \equiv x_0, \sup_{0\le t\le 1} Y_t \overset{\mathcal{D}}{=} \nu$, and $Y_1 \overset{\mathcal{D}}{=} \nu_\Delta$. For this martingale, the inequality of (5.7) is attained for all convex functions ϕ on $\mathbb{R}$. From Lemma 1.4 of [8], it may be assumed that the paths of $(Y_t)_{0\le t\le 1}$ have the following property: for each $\omega \in \Omega$, there is some $b(\omega) \in (0,1]$ for which $Y_t(\omega)$ is nondecreasing for t in $[0, b(\omega))$, and

$$\sup_{0\le t\le 1} Y_t(\omega) = Y_{b(\omega)-}(\omega) \ge Y_{b(\omega)}(\omega) = Y_u(\omega) = Y_1(\omega) \quad \text{for all } u \in [b(\omega), 1].$$

We give examples of such martingales for the examples of Section 2.

Example 5.5: (Continuation of Lemmas 2.5 and 2.6 and Examples 2.7). Let $\nu \in \mathcal{P}_0^*$, a maximal p.m., as in (2.7) and Lemma 2.6, with left inverse F_ν^{-1} defined on probability space $([0,1], \mathcal{B}([0,1]), m)$. Assume $x_0 := \inf\{z : \nu(-\infty, z] > 0\}$ is finite. The associated minimal p.m. ν_Δ is then in $\mathcal{E}$, with $\int x d\nu_\Delta = x_0$, and its left inverse $F_{\nu_\Delta}^{-1}$ is given by (2.16). Define the filtration $\{\mathcal{F}_t\}$ by $\mathcal{F}_t = \sigma\{\mathcal{B}([0,t]), (t,1]\}$ for $0 < t \le 1$ and $\mathcal{F}_0 = \{\phi, [0,1]\}$. Then the stochastic process $(Y_t)_{0\le t\le 1}$ defined by $Y_t = E(F_{\nu_\Delta}^{-1}|\mathcal{F}_t)$ is a martingale with respect to $\{\mathcal{F}_t\}$ satisfying $Y_0 \equiv \int x d\nu_\Delta = x_0$ and $Y_1 = F_{\nu_\Delta}^{-1}$. From Lemma 2.5 and the remarks thereafter, it is clear that $(Y_t)_{0\le t\le 1}$ has representation

$$(5.8)\quad Y_t(w) = F_{\nu_\Delta}^{-1}(w) \quad \text{if } 0 < w \le t \le 1, \quad \text{and} \quad = F_\nu^{-1}(t) \quad \text{if } 0 \le t < w < 1$$

for $0 < w < 1$, and that $\sup_{0\le t\le 1} Y_t = F_\nu^{-1}$. This martingale $(Y_t)_{0\le t\le 1}$ attains the inequality (5.7) simultaneously for all convex functions ϕ on $\mathbb{R}$ (provided

$\int \phi d\nu_\Delta$ exists). For example, if ν is exponentially distributed with parameter one, ν_Δ is given in Example 2.7(i), and the martingale $(Y_t)_{0\le t\le 1}$ of (5.8) is given by

$$Y_t(w) = -1 - \log(1-w) \text{ if } 0 < w \le t \le 1, \text{ and } = -\log(1-t) \text{ if } 0 \le t < w < 1.$$

In this case, examples of inequality (5.7), with $\phi(x) = |x+1|^p$ with $p > 1$ and $\phi(x) = e^{\lambda(x+1)}$ with $\lambda < 1$, are

$$E(|Y_1 + 1|^p) = \Gamma(p+1) \le E(|X_1 + 1|^p), \quad \text{and}$$
$$E(e^{\lambda Y_1}) = (1-\lambda)^{-1} e^{-\lambda} \le E(e^{\lambda X_1})$$

for all martingales $(X_t)_{0\le t\le 1}$ with $X_0 \equiv 0$ having $\sup_{0\le t\le 1} X_t$ exponentially distributed with parameter one.

Example 5.6: Continuation of Lemma 2.8 and Examples 2.9. For $\nu = (1/3)\varepsilon_0 + (1/3)\varepsilon_1 + (1/3)\varepsilon_2$, define random vector (X_0, X_1, X_2) by $(X_0, X_1, X_2) = (0,1,2)$, $(0,1,0)$ and $(0,-2,-2)$ with probability $1/3$ each; and for $\nu = (1/3)\varepsilon_0 + (1/9)\varepsilon_1 + (5/9)\varepsilon_2$, define random vector (X_0, X_1, X_2) by $(X_0, X_1, X_2) = (0,2,2), (0,1,2)$ and $(0,1,-5/2)$ and $(0,-5/2,-5/2)$ with respective probabilities $1/6, 7/18, 1/9$, and $1/3$. Define processes $(Y_t)_{0\le 1\le 1}$ by $Y_t = X_0$ if $0 \le t < 1/3, = X_1$ if $1/3 \le t < 2/3$, and $= X_2$ if $2/3 \le t \le 1$, and define filtrations $(\mathcal{F}_t)_{0\le t\le 1}$ by $\mathcal{F}_t = \{\phi, \Omega\}$ if $0 \le t < 1/3, = \sigma(X_0)$ if $1/3 \le t < 2/3$, and $= \sigma(X_0, X_1)$ if $2/3 \le t \le 1$. Then each process $(Y_t)_{0\le t\le 1}$ is a martingale with respect to the respective $(\mathcal{F}_t)$, satisfying $Y_0 \equiv 0, \sup_{0\le t\le 1} Y_t \stackrel{\mathcal{D}}{=} \nu$, and $Y_1 \stackrel{\mathcal{D}}{=} \nu_\Delta$, where p.m.'s ν_Δ were given in Examples 2.9. For the first p.m. ν, the martingale $(Y_t)_{0\le t\le 1}$ is analogous to that of (5.8); but for the second p.m., the analogy breaks down.

ACKNOWLEDGEMENT: This research was completed in part during visits of each author to the co-author's Institution. The authors wish to thank the Institut of Mathematische Statistik and Wirtschaftsmathematik of Universität Göttingen and the School of Mathematics of the Georgia Institute of Technology for kindness extended during these visits. The authors also wish to thank R. Foley and T. Hill for useful discussions concerning some of the material in section 1.

References

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

2. Y. S. Chow, H. Robbins and D. Siegmund, *Great Expectations: The Theory of Optimal Stopping*, Houghton-Mifflin, Boston, 1972.
3. Y. S. Chow and H. Teicher, *Probability Theory: Independence, Interchangeability, Martingales*, Springer-Verlag, New York, 1978.
4. K. L. Chung, *A Course in Probability Theory*, Academic Press, New York, 1974.
5. C. Dellacherie and P. A. Meyer, *Probabilities and Potential C*, North-Holland Mathematics Studies 151, North-Holland, Amsterdam, 1988.
6. L. E. Dubins and D. Gilat, *On the distribution of maxima of martingales*, Proc. Am. Math. Soc. **68** (1978), 337–338.
7. D. Gilat, *On the best order of observation in optimal stopping problems*, J. Appl. Prob. **24** (1987), 773–778.
8. R. P. Kertz and U. Rösler, *Martingales with given maxima and terminal distributions*, Isr. J. Math. **69** (1989), 173–192.
9. H. Levy, *Stochastic dominance and expected utility: survey and analysis*, preprint 1990.
10. P. A. Meyer, *Probability and Potentials*, Blaisdell, London, 1966.
11. E. Perkins, *The Cereteli–Davis solution to the H^1-embedding problem and an optimal embedding in Brownian motion*, in *Seminar on Stochastic Processes, 1985*, Birkhäuser, Boston, 1986, pp. 172–223.
12. S. I. Resnick, *Extreme Values, Regular Variation and Point Processes*, Springer-Verlag, New York, 1987.
13. R. T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, 1970.
14. D. Stoyan, *Comparison Methods for Queues and Other Stochastic Models* (D. J. Daley, ed.), Wiley, New York, 1983.
15. V. Strassen, *The existence of probability measures with given marginals*, Ann. Math. Statist. **36** (1965), 423–439.
16. D. P. van der Vecht, *Inequalities for Stopped Brownian Motion*, C.W.I. Tract 21, Mathematisch Centrum, Amsterdam, 1986.