IS EVERY SUBMARTINGALE
A CONVEX FUNCTION OF A MARTINGALE?

by David Gilat

Technical Report No. 261

February, 1976

University of Minnesota, Minneapolis*

*On leave from Tel-Aviv University, Israel.

AMS Classification: 60G45

Key Words: (super) fair process; (sub)martingale; convex function; regular conditional distribution; strategy.

IS EVERY SUBMARTINGALE A CONVEX FUNCTION OF A MARTINGALE?

by David Gilat
University of Minnesota, Minneapolis*
February, 1976

ABSTRACT

It is shown that every non-negative superfair process (in particular a non-negative submartingale) is the absolute value of a symmetric fair process (martingale). For the more general question posed in the title, the evidence is inconclusive. If however the adjective convex is omitted from the title, an affirmative answer is provided. Furthermore, transforming functions φ, such that every superfair process (submartingale) is that φ of a fair process (martingale), are shown to exist. The results are extended to continuous-parameter submartingales with right-continuous sample functions.

---

*On leave from Tel-Aviv University, Israel.

AMS Classification: 60G45

Key Words: (super) fair process; (sub)martingale; convex function; regular conditional distribution; strategy.

1. Introduction.

Let $M = (M_1, M_2, \ldots)$ be a martingale and suppose φ is a convex function such that $E|\varphi(M_n)| < \infty$ for all n. It is then an immediate and well-known consequence of Jensen's inequality ([7], p. 29) that the process $\varphi(M) = (\varphi(M_1), \varphi(M_2), \ldots)$ is a submartingale. It seems natural to inquire whether the converse is true. More precisely, attention is called to the following question: Given a submartingale $S = (S_1, S_2, \ldots)$, is there always a martingale $M = (M_1, M_2, \ldots)$ and a convex function φ, such that the process $\varphi(M) = (\varphi(M_1), \varphi(M_2), \ldots)$ has the same distribution as S, and if so, to what extent does the function φ have to depend on S? Theorem 1 provides an affirmative answer in the case when S is non-negative. Furthermore in this case the absolute-value-function, $\varphi: x \to |x|$, is shown to work for every S. Roughly speaking Theorem 1 says that every non-negative submartingale is the absolute value of a martingale. Moreover, the corresponding martingale can be chosen either to be symmetric or else to have any mean m with $|m| \leq ES_1$.

For submartingales which may assume negative values, the evidence is inconclusive so long as one insists on the convexity of the transforming function φ. If however φ is not required to be convex and one merely asks whether every submartingale is a function of a martingale, then the answer is yes. Furthermore, there are functions φ such that every submartingale is the φ of some martingale. In fact Theorem 2 shows that such a function φ can be made symmetric and resemble (see (13) of Section 2) the absolute-value-function outside an arbitrarily small neighborhood of the origin. Incidentally, Theorem 1 and 2 are shown to hold not only for (sub)martingales but for the rather more general class of (super) fair processes. The distinction being that in order for a process to qualify as a (sub)martingale it is required to have finite expectations in addition to being (super) fair.

Section 3 is devoted to the extension of Theorem 1 to continuous-parameter submartingales with right-continuous sample functions. Here the obvious excursion through the hierarchy of binary rationals is taken. The trouble is that the construction suggested in the proof of Theorem 1 does not necessarily yield a consistent family of finite-dimensional distributions. Consequently one has to resort to some weak-convergence arguments in order to establish the existence of the desired martingale. In some special cases, such as the Poisson process, it is possible to explicitly construct a martingale whose absolute values form the given submartingale. In general, however, the method of proof gives practically no insight into the nature of these martingales. It may perhaps be of interest to find out more about a martingale whose absolute value is distributed like, for example, the square of Brownian Motion.

2. Discrete-parameter processes.

As is evident from the introduction, this note is concerned with distributions of stochastic processes rather than with the processes themselves. It is therefore expedient and does no harm to identify a process with its distribution. Both terms will thus be treated as interchangeable synonyms, letting convenience dictate which is to be used in any particular statement. The distribution of a real-valued process $X = (X_1, X_2, \ldots)$ is most conveniently perceived as the sequence $\sigma = (\sigma_0, \sigma_1, \ldots)$ of its successive regular conditional distributions given the past, where of course

σ_0 is the distribution of X_1, while for each $n \geq 1$ and every n-tuple $(x_1, \ldots, x_n)$ of real numbers, $\sigma_n(x_1, \ldots, x_n)$ is a regular conditional distribution of X_{n+1} given $X_j = x_j$, $1 \leq j \leq n$. Indicative of the gambling ideas that have produced Theorem 1, σ might be called a strategic form of the process X, or simply its strategy. In fact σ is a (measurable) strategy in the sense of Dubins and Savage [4]. Plainly, two processes have the same distribution iff they admit of equal strategic forms. Henceforth a process will be identified with its strategy.

A lottery is a probability measure on the real line. If θ is a lottery with $\int|x|d\theta(x) < \infty$, write $m(\theta)$ for $\int x d\theta(x)$ = the mean of θ. A process or strategy $\sigma = (\sigma_0, \sigma_1, \ldots)$ is fair (superfair) if for each $n \geq 1$ and every n-tuple $(x_1, \ldots, x_n)$ of real numbers, $m(\sigma_n(x_1, \ldots, x_n)) = x_n$ (is finite and no less than x_n). Notice that for σ to be (super) fair, neither σ_0 nor the marginal σ-distributions of the coordinates need have a mean. Given a lottery θ, it is convenient to introduce two related lotteries $\bar{\theta}$ and $|\theta|$ defined by

$$(1) \qquad \bar{\theta}(A) = \theta(-A)$$

$$(2) \qquad |\theta| = \theta(A^+) + \bar{\theta}(A^+) - \theta(\{0\})$$

for every (Borel) subset, A, of the real line, where

$$-A = \{x: -x \epsilon A\} \quad \text{and} \quad A^+ = A \cap [0, \infty).$$

Note: $\bar{\theta}(-A) = \theta(A)$; $|\bar{\theta}| = |\theta|$; θ is symmetric iff $\bar{\theta} = \theta$; it is non-negative (i.e. $\theta[0, \infty) = 1$) iff $|\theta| = \theta$.

Similarly, a process or a strategy $\sigma = (\sigma_0, \sigma_1, \ldots)$ is non-negative if $|\sigma_0| = \sigma_0$ and if for each $n \geq 1$ and every n-tuple $(x_1, \ldots, x_n)$ of non-negative numbers, $|\sigma_n(x_1, \ldots, x_n)| = \sigma_n(x_1, \ldots, x_n)$; σ is symmetric provided $\bar{\sigma}_0 = \sigma_0$ and $\bar{\sigma}_n(-x_1, \ldots, -x_n) = \sigma_n(x_1, \ldots, x_n)$ for all

$n \geq 1$ and all n-tuples $(x_1, \ldots, x_n)$ of real numbers. These conditions on σ are of course equivalent to the corresponding conditions on the coordinate-process $X = (X_1, X_2, \ldots)$ induced by σ; i.e. σ is non-negative iff so is X; σ is symmetric iff X and $-X = (-X_1, -X_2, \ldots)$ have the same distribution.

Theorem 1. Suppose $\sigma = (\sigma_0, \sigma_1, \ldots)$ is a non-negative superfair process. Then there is a symmetric fair process $\mu = (\mu_0, \mu_1, \ldots)$, such that

(*) $|\mu_0| = \sigma_0$ and $|\mu_n(x_1, \ldots, x_n)| = \sigma_n(|x_1|, \ldots, |x_n|)$,

for all $n \geq 1$ and all n-tuples $(x_1, \ldots, x_n)$ of real numbers.

Furthermore, if σ_0 has a finite mean, then for every m with $|m| \leq m(\sigma_0)$, there is a fair process μ with $m(\mu_0) = m$ for which (*) holds.

The key to the construction of μ is a mapping α which associates with every pair (θ, x), where θ is a lottery with a finite mean and x is a real number with $|x| \leq |m(\theta)|$, another lottery, $\alpha(\theta, x)$, defined as the unique convex combination of θ and $\overline{\theta}$ whose mean is x. Formally,

$$(3) \qquad \alpha(\theta, x) = \begin{cases} \dfrac{m(\theta) + x}{2m(\theta)}\,\theta + \dfrac{m(\theta) - x}{2m(\theta)}\,\overline{\theta}, & m(\theta) \neq 0 \\ \tfrac{1}{2}\theta + \tfrac{1}{2}\overline{\theta}, & m(\theta) = 0 \end{cases} \qquad |x| \leq |m(\theta)|.$$

Some easily verifiable properties of the mapping α are listed here for later reference. All pairs $(\cdot, \cdot)$ occurring in the list, (4) through (10), are assumed to be in the domain of definition of α.

(4) $m(\alpha(\cdot, x)) = x.$

(α was aimed at these two properties.)

(5) $|\alpha(\theta, \cdot)| = |\theta|.$

(6) $\bar{\alpha}(\cdot, x) = \alpha(\cdot, -x)$.

(7) $\alpha(\bar{\theta}, \cdot) = \alpha(\theta, \cdot)$.

(8) $\alpha(\alpha(\cdot, x), y) = \alpha(\cdot, y)$.

(9) $\alpha(\cdot, \lambda x + (1 - \lambda)y) = \lambda\alpha(\cdot, x) + (1 - \lambda)\alpha(\cdot, y), \quad 0 \leq \lambda \leq 1$.

(10) $\alpha(\lambda\theta_1 + (1 - \lambda)\theta_2, \cdot) = \lambda\alpha(\theta_1, \cdot) + (1 - \lambda)\alpha(\theta_2, \cdot)$, provided $m(\theta_1) = m(\theta_2)$.

Proof of Theorem 1. The construction of μ is facilitated essentially by (4) and (5), thus. If $m(\sigma_0)$ is finite and it is desired that μ_0 have a prescribed mean m with $|m| \leq m(\sigma_0)$, set $\mu_0 = \alpha(\sigma_0, m)$; otherwise pick any $\lambda\epsilon[0, 1]$ and take $\mu_0 = \lambda\sigma_0 + (1 - \lambda)\bar{\sigma}_0$. In both instances $|\mu_0| = \sigma_0$; μ_0 is symmetric provided $m = 0$ in the first case, and when $\lambda = \frac{1}{2}$ in the second. The construction of $(\mu_1, \mu_2, \ldots)$ proceeds with no regard to the choice of μ_0. For $n \geq 1$, μ_n is defined simply by

(11) $$\mu_n(x_1, \ldots, x_n) = \alpha(\sigma_n(|x_1|, \ldots, |x_n|), x_n),$$

where $(x_1, \ldots, x_n)$ is any n-tuple of real numbers. Observe that $|x_n| \leq m(\sigma_n|x_1|, \ldots, |x_n|)$ because σ is superfair and thus μ_n is well defined. Clearly, (4) implies that the process $\mu = (\mu_0, \mu_1, \ldots)$ is fair, whereas that μ satisfies (*) is an immediate consequence of (5).

The issue of symmetry is settled by means of (6). As has been demonstrated, μ_0 can always be made symmetric simply by taking it to be $\frac{1}{2}\sigma_0 + \frac{1}{2}\bar{\sigma}_0$. Having done so, the entire process μ, as constructed, is necessarily symmetric because

(12) $$\bar{\mu}_n(-x_1, \ldots, -x_n) = \bar{\alpha}(\sigma_n(|x_1|, \ldots, |x_n|), -x_n)$$
$$= \alpha(\sigma_n(|x_1|, \ldots, |x_n|), x_n) = \mu_n(x_1, \ldots, x_n),$$

where the first and last equalities are the definition, (11), of μ_n, while the middle equality follows by an application of (6). Equality of the extreme sides of (12) for all $n \geq 1$ and all $(x_1, \ldots, x_n)$, together with the symmetry of μ_0, is precisely the symmetry of the entire process μ. Incidentally, (12) alone can naturally be interpreted as conditional symmetry given the initial state of the process. It is thus evident that all the fair processes μ obtained here are, in this sense, conditionally symmetric. Another pleasant feature of the method is that if σ is fair to begin with and one chooses $\mu_0 = \sigma_0$, then the construction yields $\mu = \sigma$.

For the rest of this section it is convenient to switch back to the traditional language of sequences of random variables. A (sub)martingale is of course a (super) fair process $X = (X_1, X_2, \ldots)$ such that $E|X_n| < \infty$ for all n.

Corollary 1. *Given a non-negative submartingale* $S = (S_1, S_2, \ldots)$, *then for any real number* m *with* $|m| \leq ES_1$, *there is a (conditionally symmetric) martingale* $M = (M_1, M_2, \ldots)$ *with mean* m *such that, the process* $|M| = (|M_1|, |M_2|, \ldots)$ *has the same distribution as* S. *If* $m = 0$, *then* M *can be made symmetric*.

Proof. Translate from strategic to random-variable terminology and interpret Theorem 1 accordingly.

Proceed now to general submartingales. For each $\epsilon > 0$, introduce the function $\varphi = \varphi_\epsilon$ defined by

$$(13) \qquad \begin{aligned} \varphi(x) &= |x| - \epsilon, && |x| \geq \epsilon \\ &= \epsilon(\ln|x| - \ln\epsilon), && 0 < |x| < \epsilon. \end{aligned}$$

Theorem 2. *Suppose* $S = (S_1, S_2, \ldots)$ *is any superfair process (submartingale). Then for every* $\varepsilon > 0$, *there is a fair process (martingale)* $M_\varepsilon = M = (M_1, M_2, \ldots)$ *such that the process* $\varphi_\varepsilon(M) = (\varphi_\varepsilon(M_1), \varphi_\varepsilon(M_2), \ldots)$ *has the same distribution as* S.

Proof. Given $\varepsilon > 0$, consider the function $\Psi = \Psi_\varepsilon$ defined by

$$(14) \qquad \Psi(x) = \begin{cases} \varepsilon e^{x/\varepsilon} & x < 0 \\ x + \varepsilon & x \geq 0. \end{cases}$$

Plainly, Ψ is positive, convex and increasing. Therefore $\Psi(S) = (\Psi(S_1), \Psi(S_2), \ldots)$ is a positive superfair process. Apply Theorem 1 to $\Psi(S)$ to obtain a fair process M, such that $|M|$ has the same distribution as $\Psi(S)$. Next apply Ψ^{-1} to both of these processes to obtain that $\Psi^{-1}(|M|)$ and S have the same distribution. Finally, observe that $\Psi_\varepsilon^{-1}(|\cdot|) = \varphi_\varepsilon(\cdot)$. If S is a submartingale (i.e. superfair with $E|S_n| < \infty$), it is easy to check that so is $\Psi(S)$, therefore in this case M is indeed a martingale and not merely a fair process. The proof is thus complete.

3. Continuous-parameter processes.

Theorem 3. *Let* $Z = \{Z_t, t \geq 0\}$ *be a non-negative submartingale, almost all of whose sample-functions are right-continuous. Then there exists a martingale* $Y = \{Y_t, t \geq 0\}$ *such that, the process* $Y = \{|Y_t|, t \geq 0\}$ *has the same distribution as* Z. *Furthermore,* Y *can be chosen either to be symmetric or else to have any mean* m *with* $|m| \leq EZ_0$.

Let T be a countable dense subset of $[0, \infty)$, such as for example the set of binary rationals. The major step in the proof of Theorem 3 consists of demonstrating the existence of a martingale $\{Y_t, t\epsilon T\}$ for which the distribution of $\{|Y_t|, t\epsilon T\}$ is the same as that of $\{Z_t, t\epsilon T\}$. The existence of such a martingale is the content of Theorem 3*. Before Theorem 3* can be conveniently stated, some handy notation is needed.

Let T be any countable subset, of $[0, \infty)$ (think of T as being the set of rationals in $[0, \infty)$). Let Ω be the set of all real-valued functions on T. For $t\epsilon T$ and $\omega\epsilon\Omega$, let $X_t(\omega) = \omega(t)$. Take $\mathcal{B}$ to be the smallest sigma-algebra of subsets of Ω with respect to which every X_t, $t\epsilon T$, is measurable from $(\Omega, \mathcal{B})$ to the Borel real line, and, for $t\epsilon T$, $\mathcal{B}_t$ to be the sigma-algebra generated by the collection $\{X_s, s\epsilon T, s \leq t\}$. A probability measure σ on $\mathcal{B}$ turns the coordinate-process, $\{X_t, t\epsilon T\}$, of Ω into a real-valued stochastic-process whose paths are points in ω. Denote by E_σ expectations as well as conditional expectations with respect to σ. Identfying a process with its distribution, refer to σ as being the process itself. Say that σ is non-negative, if for each $t \geq 0$, $\sigma\{\omega: X_t(\omega) \geq 0\} = 1$; that σ is a (sub)martingale, if $E_\sigma|X_t| < \infty$ for all $t\epsilon T$, and $E_\sigma(X_t|\mathcal{B}_s)(\geq) = X_s$ for all $s \leq t$, s and t both in T.

Theorem 3* *Let σ be a non-negative submartingale on $(\Omega, \mathcal{B})$. Then there exists a martingale μ on $(\Omega, \mathcal{B})$ such that the μ-distribution of $\{|X_t|, t\epsilon T\}$ is σ. Furthermore, given any real number m with $|m| \leq \inf_{t\epsilon T} E_\sigma X_t$, μ can be chosen to have mean m (i.e. $E_\mu X_t = m$, all $t\epsilon T$). When $m = 0$, the μ obtained is symmetric.*

Proof. Given a subset, S, of T, let $\mathcal{B}(S)$ be the sub sigma-algebra of $\mathcal{B}$ generated by $\{X_t, t\epsilon S\}$. $\mathcal{B}(S)$ is of course isomorphic to the product sigma-algebra on the set of all functions from S to the real line R, R^S. For each t in S let $\mathcal{B}_t(S)$ be the further sub sigma-algebra of $\mathcal{B}(S)$ generated by $\{X_s, s\in S, s \leq t\}$. Let σ_S denote the restriction of σ to $\mathcal{B}(S)$. Clearly, under σ_S, the process $\{X_t, t\in S\}$ forms a non-negative sub-martingale with respect to its intrinsic sigma-algebras $\{\mathcal{B}_t(S), t\in S\}$. Suppose now that S is a finite subset of T. Theorem 1, then, applies to obtain a probability measure μ_S on $(\Omega, \mathcal{B}(S))$ for which the adapted process $\{(X_t, \mathcal{B}_t(S)), t\in S\}$ is a martingale with the prescribed mean m and such that, the μ_S-distribution of $\{|X_t|, t\in S\}$ is the same as the σ_S-distribution of $\{X_t, t\in S\}$. Doing so for every finite subset, S, of T, produces a system of finite-dimensional distributions, $\{\mu_S\}$, for the coordinate-process, $\{X_t, t\in T\}$, on Ω. Unfortunately, however, as pointed out in the introduction, the system $\{\mu_S\}$ is generally not consistent and therefore it does not extend as such to a measure μ on the full sigma-algebra $\mathcal{B}$. An additional argument is thus needed to obtain the desired μ.

Enumerate T and arrange it in a sequence $(t_1, t_2, \ldots)$. For $n \geq 1$, let $S(n)$ be the set $\{t_1, \ldots, t_n\}$, reordered so as to form an increasing sequence of real numbers. Abbreviate $\mathcal{B}(S(n))$ by $\mathcal{B}(n)$, $\mu_{S(n)}$ by μ_n and let μ_n^k, $1 \leq k \leq n$, be the restriction of μ_n from $\mathcal{B}(n)$ to $\mathcal{B}(k)$. Since μ_n^k is essentially a probability measure on the Borel sigma-algebra of Euclidean k-space, the standard diagonal method (see for example p. 205 of [5]) applies to obtain a subsequence, $\{n'\}$, of $\{n\}$, and for each k a sub-probability-measure μ^k on $\mathcal{B}(k)$, such that $\{\mu_{n'}^k\}$ converges weakly to μ^k. That the μ^k are proper probability measures, follows from the fact that for each fixed k, the μ_n^k-distribution of $\{|X_t|\}, t\in S(k)\}$ is $\sigma_{S(k)}$, independently of $n \geq k$. Therefore for each k, the sequence $\{\mu_n^k, n \geq 1\}$ is tight and no mass can escape in the limiting process. By the very nature of the diagonal method, $\{\mu^k\}$ forms a consistent system of finite-dimensional distributions on $\{\mathcal{B}(k)\}$; and since $S(k)$ increases to T, Kolmogorov's consistency theorem applies to obtain a probability measure μ on $(\Omega, \mathcal{B})$ whose restriction to $\mathcal{B}(S)$, for any finite subset, S, of T, is μ^S, where μ^S is of course the restriction of μ^k to $\mathcal{B}(S)$ for any k such that $S(k) \supset S$. Plainly, the μ-distribution of $\{|X_t|, t\epsilon T\}$ is σ, $E_\mu X_t = m$ for all t in T and, if $m = 0$, $\{X_t, t\epsilon T\}$ and $\{-X_t, t\epsilon T\}$ have the same μ-distributions. Also, it is not hard to argue that under μ, the adapted process $\{(X_t, \mathcal{B}_t(S)), t\epsilon S\}$ forms a martingale, for every finite $S \subset T$. That under such circumstances, the entire coordinate process $\{(X_t, \mathcal{B}_t), t\epsilon T\}$ forms a martingale is perhaps noteworthy for its own sake, especially when contrasted with an example, due to Diudonné [1], of a uniformly-integrable, countable martingale-<u>net</u> which fails to converge in the almost-sure sense. To establish this fact, let $s < t$ be two fixed elements of T. Let C be the collection of all finite sets F, such that $\{s, t\} \subset F \subset T$. For F in C, let $Y_F = E(X_t|\mathcal{B}_s(F))$. Since the pair $\{(X_s, \mathcal{B}_s(F)), (X_t, \mathcal{B}_t(F))\}$ forms a martingale, $Y_F = X_s$ a.s. for every F in C. In particular Y_F converges a.s. to X_s, as F <u>filters</u> to T. On the other hand, when C is ordered by inclusion, the process $\{(Y_F, \mathcal{B}_s(F)), F\epsilon C\}$ forms a uniformly integrable martingale-<u>net</u>. Observe that since T is countable, $\sup \{\mathcal{B}_s(F), F\epsilon C\} = \mathcal{B}_s(T) = \mathcal{B}_s$. So, by Helms [6], Y_F converges in L_1 to $E(X_t|\mathcal{B}_s)$ as F filters to T. Of course, the L_1-limit of Y_F has to agree a.s. with its almost-sure limit and consequently $E(X_t|\mathcal{B}_s) = X_s$ a.s.

The proof of Theorem 3* is thus complete.

Theorem 3 follows from 3* by standard martingale arguments, such as can be found, for example, in Chapter 6 of Meyer [7].

Acknowledgement. Steven Orey's initial skepticism and later interest regarding the validity of Theorems 1 and 3, have lead to their proofs. The idea of Theorem 2 occurred to me after a conversation with Joe Eaton.

References.

[1] J. Diudonné (1950), Sur un théorème de Jensen, *Fund. Math.* Vol. 37, 242-248.

[2] L. E. Dubins (1968), On a theorem of Skorohod, *Ann. Math. Stat.* Vol. 39, No. 6, 2094-2097.

[3] L. E. Dubins and G. Schwarz (1965), On extremal Martingale distributions, *Proc. of the Fifth Berkely Symp. on Math. Stat. and Prob.* Vol. II, Part 1, 295-299.

[4] L. E. Dubins and L. J. Savage, *How to Gamble If You Must*, McGraw-Hill, New York, 1965.

[5] D. Freedman, *Brownian Motion and Diffusion*, Holden-Day, Inc., 1971.

[6] L. L. Helms (1958), Mean convergence of martingales, *Tran. Am. Math. Soc.* Vol. 87, 439-446.

[7] P. A. Meyer, *Probability and Potentials*, Blaisdell Publishing Co., 1966.

David Gilat
Department of Theoretical Statistics
270 Vincent Hall
University of Minnesota
Minneapolis, Minnesota 55455