MOMENT SPACES FOR IFR DISTRIBUTIONS, APPLICATIONS AND RELATED MATERIAL

by

Sujit K. Basu* and Gordon Simons**
University of North Carolina, Dept. of Statistics

April 1982

*Permanently at the Indian Institute of Management, Calcutta, India.

**His research was supported by The National Science Foundation Grant MCS81-00748.

IN HONOR OF

Professor Norman L. Johnson

ABSTRACT

Some useful topological properties of the moment space of the family of distributions with increasing failure rates are established. These results are exploited to obtain tight inequalities for moments of such distributions and also to derive some interesting facts concerning weak convergence within this family.

Key Words and Phrases: IFR distributions, total positivity, weak convergence, moment inequalities, extremal distributions, exponential distribution.

AMS 1980 Subject Classifications. Primary 62N05; Secondary 60E15, 60F22

1. INTRODUCTION AND SUMMARY

An important variant of the "problem of moments" (referred to as the "reduced moment problem" by Shohat and Tamarkin (1943)) is the problem of describing the "moment space" M_n of points $c=(c_1,\ldots,c_n)$ in n-dimensional euclidean space whose components are the first n moments (or n distinct moments) of a distribution function (d.f.) which is a member of some specified family F of d.f.'s. Besides the intrinsic interest mathematicians have found in this problem and its extensions, the problem arises naturally within the larger context of finding Chebyshev inequalities when various moments are given. Moreover, the study of moment spaces leads naturally to the establishment of tight moment inequalities. For instance, if $c \in M_n$ then typically $(c_1,\ldots,c_n,x) \in M_{n+1}$ for all x in some closed interval $[a,b]$. Thus if X is a random variable (with d.f. in F) whose first n moments equal $c_1,\ldots,c_n$, respectively, then its (n+1)-st moment μ_{n+1} satisfies the tight inequalities $a \le \mu_{n+1} \le b$. The sizes of a and b (and therefore the usefulness of these inequalities) depend upon the family F with which one is working. Mallows (1956) has argued that the "smoothness" of the d.f.'s in F affects the quality of the Chebyshev inequalities that can be obtained. Presumably this applies to moment inequalities as well.

Examples of families F that have been considered are:

(a) *families defined on specified intervals in* $\mathbb{R}$ Descriptions of M_n, expressed in terms of quadratic forms or Hankel determinants, can be found for finite intervals (Shohat and Tamarkin (1943, p.77)), for semi-infinite intervals (Karlin and Studden (1966, p.173)), and $\mathbb{R}$ (Ahiezer and Krein (1962, pp.3,4));

(b) *absolutely continuous families defined on* $\mathbb{R}$ *whose densities have a specified upper bound* (Shohat and Tamarkin (1943, pp. 82,83));

(c) *various families of unimodal d.f.'s* (Johnson and Rogers (1951), Mallows (1956));

(d) *families of d.f.'s whose densities are in various Pólya frequency classes* (Karlin, Proschan and Barlow (1961)).

The present paper is concerned with the family of d.f.'s which possess the increasing failure rate (IFR) property. (The importance of this family has been well-documented, and no further justification will be attempted here.) The theory we develop for M_n is geometric in character, and, in many ways, it closely parallels the theories developed by Krein (1959), Karlin and Shapley (1953), and Karlin and Studden (1966), for families of finite measures. (A clear, succinct description of Krein's results has been given by Mallows (1963).) In particular, their theories and ours both focus attention, for each n, on two families of "extremal distributions" (or measures), which are explicitly described in geometric terms. However, whereas their approaches rely heavily on the mathematics of convex cones, ours can not. This is because the families F (of finite measures) with which they are concerned are closed under the formation of convex mixtures, while the family of IFR distributions is not. Our approach draws, simply, upon the mathematics of algebraic topology, and (as with their approaches) on the ideas associated with the concept of total positivity. We have not tried to determine whether our methods are suitable for their settings.

There is a long history associated with the problem of moments. The early history is well documented in the extensive survey provided by Shohat and Tamarkin (1943). Much of the recent history has been described by Karlin and Studden (1966), where a detailed account of Chebyshev systems is given. (These systems are sequences of functions which have certain important properties in common with the sequence $1, t, t^2, \ldots$.) When they are used to define "generalized moments", most of the classical results, associated with the problem of moments, generalize. Undoubtedly such generalizations could be made of our results. However, we are unaware of any that are likely to be of interest to statisticians, or to those concerned with the area of reliability.

Section 2 introduces the extremal distributions with which we must work, and some notation. Section 3 develops the moment-space theory for the class of IFR distributions, and Section 4 applies this theory to obtain tight moment inequalities. Some moment inequalities are known for IFR distributions; these reappear as a consequence of our theory, together with new inequalities. Section 5 discusses the subject of weak convergence (convergence in law) for the IFR family of distributions. In particular, we apply the theory of Section 3 to obtain a generalization of a result due to Obretenov (1977). In addition, we show that weak convergence within the IFR family of distributions in essentially equivalent to the usually much stronger convergence in total variation.

2. THE EXTREMAL DISTRIBUTIONS

Let $\mathcal{F}$ denote the family of distribution functions (d.f.'s) which are IFR on $[0,\infty)$; a d.f. F on $[0,\infty)$ belongs to $\mathcal{F}$ iff $F(0) = 0$ and $-\log \overline{F}$ is a convex function on the interval $I_F = \{x \geq 0; F(x) < 1\}$ where $\overline{F}(x)$ denotes $1-F(x)$. It follows from the convexity that F is absolutely continuous on I_F. However, when I_F is bounded, F can be discontinuous at the right endpoint of I_F.

Various subfamilies of $\mathcal{F}$ must be described which play essential roles in the theory to follow. These families are most easily described geometrically by means of "graphs". For any d.f. F on $[0,\infty)$, let $\mathcal{L}(F) = \{(x,y): \overline{F}(x) \leq e^{-y} \leq \overline{F}(x-)\}$, called (here) the *graph* of F. Membership in $\mathcal{F}$ can be characterized in terms of graphs as follows: $F \epsilon \mathcal{F}$ iff $\mathcal{L}(F)$ is a continuous, nondecreasing, convex curve in the region $[0,\infty) \times [0,\infty)$ which originates at (0,0), is unbounded in y, is bounded or unbounded in x(as I_F is bounded or unbounded), and may or may not have an infinite slope (at the right endpoint of I_F) when bounded in x. Typical graphs are shown in Figure 1.

Let $\mathcal{F}^*$ denote the subfamily of $F \epsilon \mathcal{F}$ for which $\mathcal{L}(F)$ is a polygonal line with a finite number of segments. Partition $\mathcal{F}^*$ into two subfamilies $\mathcal{G}$ and $\mathcal{H}$ on the basis of whether I_F is bounded or unbounded, respectively, and then partition $\mathcal{G}$ and $\mathcal{H}$ as $\mathcal{G} = \sum_{n=1}^{\infty} \mathcal{G}_n$ and $\mathcal{H} = \sum_{n=1}^{\infty} \mathcal{H}_n$ as follows: A d.f. $F \epsilon \mathcal{G}$ belongs to $\mathcal{G}_n$ if the polygonal line $\mathcal{L}(F)$ possesses $[\frac{n+1}{2}] + 1$ segments (where [a] denotes the integer-part of a), with the left-most segment horizontal or of positive slope as n is odd or even, and with a vertical right-most segment. A d.f. $F \epsilon \mathcal{H}$ belongs to $\mathcal{H}_n$ if the polygonal line $\mathcal{L}(F)$ possesses $[\frac{n}{2}] + 1$ segments, with left-most segment horizontal or of positive slope as n is even or odd (the reverse of $\mathcal{G}_n$) and right-most segment necessarily of finite, positive slope. Various examples are illustrated in

Figures 2a, 2b and 2c. The families G_1 and H_1 refer to d.f.'s of random variables (r.v.'s) which are degenerate and exponential respectively. The families G_2 and H_2 refer to the d.f.'s or r.v.'s of the forms min(X,c) and X + c, respectively, where X is an exponential r.v. and $c > 0$ is a constant. Inductive descriptions of the families G_n and H_n can be given but will not be needed.

An important fact about the families G_n and H_n is that its members are uniquely determined by n positive parameter $x_1,\ldots,x_n$, required to specify the nodes and slopes of their graphs. For instance, for $F \in G_4$, the location of the two nodes are specified by the values of x_1 and $x_1 + x_3$, while the two slopes are specified by x_2 and $x_2 + x_4$. See Figure 3. It is easily checked that this parameterization defines a homeomorphic mapping (one-to-one and bicontinuous) from $\mathbb{R}_+^4$ (where $\mathbb{R}_+ = (0,\infty)$) *onto* G_4. (Here, continuity is defined using the topology associated with euclidean distance for $\mathbb{R}_+^4$, and the topology associated with weak convergence [i.e., convergence in law] for F.) The general situation is described in Proposition 10 below.

The families G_n and H_n, for small values of n, have figured in the works of Karlin, Proschan and Barlow (1961), and of Barlow and Marshall (1964a and b, 1967). In the latter papers, these d.f.'s are referred to as "extremal distributions", a terminology which C. L. Mallows attributes to M.G. Krein.

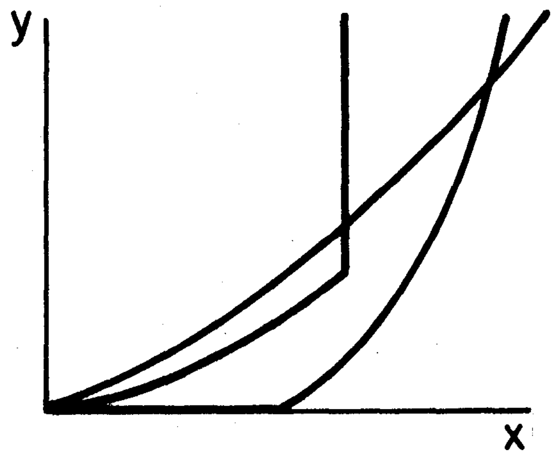

Figure 1

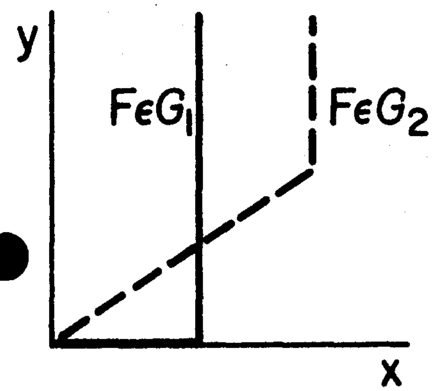

Figure 2a

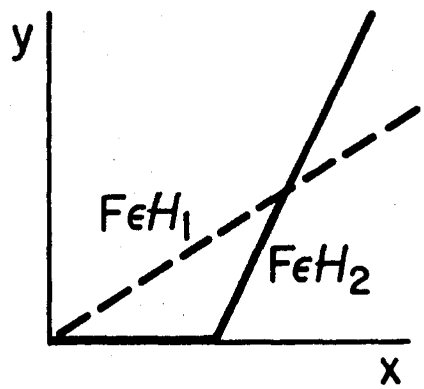

Figure 2b

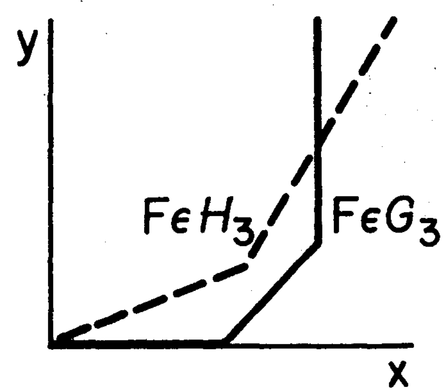

Figure 2c

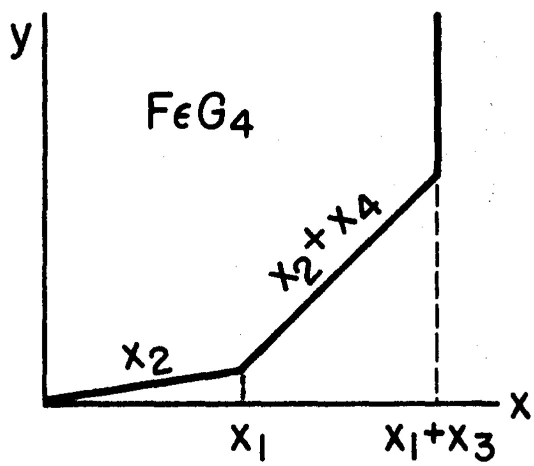

Figure 3

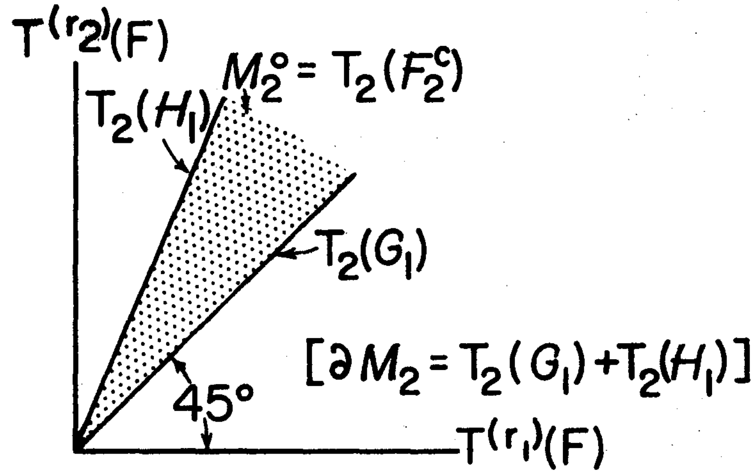

Figure 4

3. MOMENT SPACES

Here we assume that the material in Section 2 is known and applies. Let $T^{(r)}: \mathcal{F} \to \mathbb{R}_+$ $(=(0,\infty))$ denote the functional defined by $T^{(r)}(F) = (\int_0^\infty x^r dF(x))^{1/r}$, $r > 0$, and for a strictly increasing sequence $r_1, r_2, \ldots$ of positive numbers (hereafter fixed), let $T_n: \mathcal{F} \to \mathbb{R}_+^n$ denote the functional defined by $T_n(F) = (T^{(r_1)}(F), \ldots, T^{(r_n)}(F))$, $n \geq 1$. The focus of attention in this section is upon the *moment spaces* $M_n = T_n(\mathcal{F})$, $n \geq 1$. We shall describe their shape (conic, Theorem 1), their boundaries (∂M_n, Theorem 2), and their interiors (M_n^o, Theorem 3). Theorems 4 and 5 summarize the main facts from which one can establish tight moment inequalities. We shall first describe the main results, which are expressed as the theorems, then the lesser results as propositions, and then some working facts in the forms of lemmas. The proofs of the main results will be given toward the end of the section, after the propositions have been established. Throughout, the topology of M_n is just the standard euclidean topology inside $\mathbb{R}_+^n$, and the topology of $\mathcal{F}$ is that associated with weak convergence (convergence in law).

Theorem 1. *M_n is a closed cone,* $n \geq 1$.

This result would be false if we had defined $T^{(r)}$ without an r-th root. The space M_1 is simply $\mathbb{R}_+$, and the space M_2 is a wedge in $\mathbb{R}_+^2$ with lower slope 1 (arising from the degenerate distributions) and upper slope $\{\Gamma(r_2+1)\}^{1/r_2}/\{\Gamma(r_1+1)\}^{1/r_1}$ (arising from the exponential distributions). See Figure 4. Both of these spaces are *convex* cones. Unfortunately,

M_3 is not a convex set (a fact that can be shown by using the description of the upper boundary of M_3, namely $T_3(H_2)$, given in Theorem 4).

Let $F_n = \sum_{i=1}^{n-1} \{G_i + H_i\}$ and $F_n^c = F - F_n$, the complement of F_n, $n \geq 1$.

<u>Theorem 2 (the boundary of M_n).</u> $\partial M_n = T_n(F_n) = \sum_{i=1}^{n-1} \{T_n(G_i) + T_n(H_i)\}$, $n \geq 2$. (∂M_1 *is empty.*) *Moreover, for* $n \geq 2$, *the map* $T_n : F_n \to \mathbb{R}_+^n$ *is one-to-one so that each boundary point of* M_n *arises from exactly one extremal distribution.*

The boundary can be described as follows: It consists of two rays $T_n(G_1)$ and $T_n(H_1)$; two surfaces $T_n(G_2)$ and $T_n(H_2)$ joining these rays; ...; and finally two (n-1)-dimensional surfaces $T_n(G_{n-1})$ and $T_n(H_{n-1})$ joining $T_n(G_{n-2})$ and $T_n(H_{n-2})$. (A similar description of boundary sets has been given by Mallows (1963) when discussing some moments spaces studied by Krein (1959).) Thus the bulk of the boundary of M_n is generated by the extremal distributions in G_{n-1} and H_{n-1}. We shall not attempt to document all of this mathematically, but the dominant roles of G_{n-1} and H_{n-1}, in generating the boundary of M_n, are revealed in Theorem 4 below.

<u>Theorem 3 (the interior of M_n).</u> $M_n^\circ = T_n(F_n^c) = T_n(G_n) = T_n(H_n)$, $n \geq 1$. ($M_1^\circ = T_n(F) = \mathbb{R}_+$.) *Moreover, for* $n \geq 1$, *the maps* $T_n : G_n \to \mathbb{R}_+^n$ *and* $T_n : H_n \to \mathbb{R}_+^n$ *are one-to-one, so that associated with each* $F \in F_n^c$ *there are unique extremal distribution functions* $G \in G_n$ *and* $H \in H_n$ *for which* $T_n(G) = T_n(H) = T_n(F)$.

It would be consistent with Krein's (1959) terminology to refer to G and H as the "lower and upper principal representatives" of the point $T_n(F)$

in M_n°. The appropriateness of using the adjectives "lower" and "upper" is made clear in Theorem 5 below (when $r = r_{n+1}$).

Theorem 4. *For* $n \geq 2$ *and each* $F \in F_{n-1}^c$, *there are unique extremal distribution functions* $G \in G_{n-1}$ *and* $H \in H_{n-1}$ *for which*

$$T_{n-1}(G) = T_{n-1}(H) = T_{n-1}(F) \text{ and } T^{(r_n)}(G) \leq T^{(r_n)}(F) \leq T^{(r_n)}(H). \tag{1}$$

The first inequality is strict unless $F = G$, *and the second is strict unless* $F = H$. *Thus both are strict if and only if* $F \in F_n^c$, *i.e., if and only if* $T_n(F) \in M_n^\circ$.

Let $G_n^+ = G_n + \sum_{i=1}^{n-1} (G_i + H_i)$ and $H_n^+ = H_n + \sum_{i=1}^{n-1} (G_i + H_i)$

denote the closures within F of the extremal families G_n and H_n, respectively. (See Proposition 1 below.) These families of extremal distribution functions appear in the following variant of Theorem 4.

Theorem 5. *For each* $n \geq 1$ *and* $F \in F$, *there are unique extremal distribution functions* $G \in G_n^+$ *and* $H \in H_n^+$ *for which* $T_n(G) = T_n(H) = T_n(F)$. *Moreover, for each* $r > 0$,

$$(-1)^k T^{(r)}(G) \leq (-1)^k T^{(r)}(F) \leq (-1)^k T^{(r)}(H) , \tag{2}$$

where k *equals the number of* $r_i \geq r$, $i=1, \ldots, n$. *In particular,*

$$T^{(r_{n+1})}(G) \leq T^{(r_{n+1})}(F) \leq T^{(r_{n+1})}(H) . \tag{3}$$

Proposition 1. *Under the topology of weak convergence (convergence in law) within* F, *the closures of the classes* G_n *and* H_n *become* $G_n^+ = G_n + F_n$ *and* $H_n^+ = H_n + F_n$, *respectively, where* $F_n = \sum_{i=1}^{n-1} (G_i + H_i)$ (F_1 *is empty*), $n \geq 1$.

Proof. This is geometrically obvious; refer to the graphs $L(G)$, $G \in G_n$, and $L(H)$, $H \in H_n$, and consider the possible "limits" of these graphs. □

Proposition 2. *If* F_1 *and* F_2 *are distinct distribution functions in* $\mathcal{F}$, *then the difference* $\overline{F}_2 - \overline{F}_1$, *changes sign at least once for each root of the function* k *defined by* $k(r) = \int_0^\infty x^r d(F_2(x) - F_1(x))$, $r > 0$. *If the numbers of sign changes of* $\overline{F}_2 - \overline{F}_1$ *and roots of* k *are finite and equal, then* $\overline{F}_2 - \overline{F}_1$ *and* k *change sign an equal number of times, and they exhibit the same arrangement of signs.*

Proof. This result is essentially known (c.f., Barlow and Proschan (1975, p.120)); it is not important that F_1 and F_2 belong to $\mathcal{F}$. An elementary proof is to write k as $k(r) = \int_0^\infty K(x,r)[\overline{F}_2(x) - \overline{F}_{\cdot}(x)]dx$ where $K(x,r) = rx^{r-1}$, $x > 0$, $r > 0$, to observe that K is strictly totally positive, and to apply Karlin's (1968, p.233) Theorem 3.1. □

Remark. Clearly when $\overline{F}_2 - \overline{F}_1$ and k change signs an equal number of times, k must change sign at each of its roots; tangencies to the x-axis, without crossing, are ruled out.

Proposition 3. *For any distribution functions* $G_1, G_2 \in \mathcal{G}_n$ *and* $H_1, H_2 \in \mathcal{H}_n$, *the differences* $\overline{G}_2 - \overline{G}_1$ *and* $\overline{H}_2 - \overline{H}_1$ *each can have at most* $n-1$ *sign changes. Thus if* $T_n(G_1) = T_n(G_2)$ *then* $G_1 = G_2$, *and if* $T_n(H_1) = T_n(H_2)$, *then* $H_1 = H_2$. $(n \geq 1)$

Proof. The first assertion is geometrically obvious; refer to the graphs of the d.f.'s G_1, G_2, H_1 and H_2. The second assertion follows immediately from Proposition 2, since, for example, the equality $T_n(G_1) = T_n(G_2)$ implies that G_1 and G_2 have at least n moments in common. □

Proposition 4. *The mappings* $T_n: \mathcal{G}_m \to \mathbb{R}_+^n$ *and* $T_n: \mathcal{H}_m \to \mathbb{R}_+^n$ *are one-to-one for* $n \geq m \geq 1$.

Proof. This follows immediately from the last assertions in Proportion 3. □

Proposition 5. *For any distribution functions* $F \in \mathcal{F}$, $G \in \mathcal{G}_n$, $H \in \mathcal{H}_n$, *the differences* $\overline{F}-\overline{G}$ *and* $\overline{H}-\overline{F}$ *each can have at most* n *sign changes. Whenever* n *sign changes occur, the last change is always from a negative to a positive sign. If* $T_{n+1}(F) = T_{n+1}(G)$, *then* $F = G$, *and if* $T_{n+1}(F) = T_{n+1}(H)$, *then* $F = H$. $(n \geq 1)$.

Proof. The first two assertions are geometrically obvious (using graphs), and the third follows immediately from Proposition 2. □

Proposition 6. *If* $F \in \mathcal{F}$, $G \in \mathcal{G}_{n-1}$, $T_{n-1}(F) = T_{n-1}(G)$ *and* $F \neq G$, *then* $F \notin \mathcal{G}_{n-1}^+$ *and* $T^{(r_n)}(F) > T^{(r_n)}(G)$. *If* $F \in \mathcal{F}$, $H \in \mathcal{H}_{n-1}$, $T_{n-1}(F) = T_{n-1}(H)$ *and* $F \neq H$, *then* $F \notin \mathcal{H}_{n-1}^+$ *and* $T^{(r_n)}(F) < T^{(r_n)}(H)$. $(n \geq 2)$

Proof. Suppose $F \in \mathcal{F}$, $G \in \mathcal{G}_{n-1}$, $T_{n-1}(F) = T_{n-1}(G)$ and $F \neq G$. If F were in $\mathcal{G}_{n-1}^+$, then (see Proposition 1) it would have to be in either $\mathcal{G}_{n-1}$ or $\mathcal{F}_{n-1} = \sum_{i=1}^{n-2} (\mathcal{G}_i + \mathcal{H}_i)$. The former case is ruled out by Proposition 3 since $T_{n-1}(F) = T_{n-1}(G)$ and $F \neq G$. The latter case is ruled out by Proposition 5 since $T_{i+1}(G) = T_{i+1}(F)$ and $G \neq F$ when $1 \leq i \leq n-2$. (The roles of F and n appearing in Proposition 5 are assumed here by G and i, respectively; the role of G or H appearing in Proposition 5 is assumed here by F, depending upon whether $F \in \mathcal{F}_{n-1}$ were to occur with $F \in \mathcal{G}_i$ or with $F \in \mathcal{H}_i$.) Thus $F \notin \mathcal{G}_{n-1}^+$.

Since $F \in \mathcal{F}$ and $G \in \mathcal{G}_{n-1}$, the difference $\overline{F}-\overline{G}$ can have at most n-1 sign changes (see Proposition 5). Since, in addition, $T_{n-1}(F) = T_{n-1}(G)$, F and G must have at least n-1 moments in common. Consequently $\overline{F}-\overline{G}$ has exactly n-1 sign changes (see Proposition 2, F and G are distinct), and the last change is from a negative to a positive sign (see Proposition 5 again). It follows that the number of sign changes of $\overline{F}-\overline{G}$ and the number of roots of the function k defined by $k(r) = \int_0^\infty x^r d(F(x)-G(x))$, $r > 0$, are finite and equal (=n-1), and the two function $\overline{F}-\overline{G}$ and k exhibit the same arrangement of signs (see Proposition 2 again). Since the sign of $\overline{F}(x)-\overline{G}(x)$ is (strictly) positive when x is large, the sign of k(r) is strictly positive when $r > r_{n-1}$. Thus $T^{(r_n)}(F) > T^{(r_n)}(G)$. The same kind of reasoning applies to F and H. □

Proposition 7. *The map* $T_n: \mathcal{F} \to \mathbb{R}_+^n$ *is continuous for each* $n \geq 1$.

Proof. Suppose $F_m \to F$ as $m \to \infty$ (in the sense of convergence in law) for $F, F_m \in \mathcal{F}$ $(m \geq 1)$. Let x_0 be point of continuity of F for which $\overline{F}(x_0) < 1$, so that $\overline{F}_m(x_0) \leq e^{-\theta}$, $m \geq m_0$, for some sufficiently small $\theta > 0$ and sufficiently large $m_0 \geq 1$. Since F_m is IFR, $-\log \overline{F}_m$ is convex on $I_{F_m} = \{x \geq 0 : F_m(x) < 1\}$, and, consequently, $-x^{-1}\log \overline{F}_m(x)$ is nondecreasing in I_{F_m} when $x > 0$. It follows that

$$\overline{F}_m(x) \leq (\overline{F}_m(x_0))^{x/x_0} \leq e^{-x\theta/x_0}, \quad x \geq x_0,\ m \geq m_0 \quad . \qquad (4)$$

Using this, together with the dominated convergence theorem, one can easily show that for each $r > 0$, $r\int_0^\infty x^{r-1}\overline{F}_m(x)dx \to r\int_0^\infty x^{r-1}\overline{F}(x)dx$ as $m \to \infty$, from which the proposition directly follows. □

Remark. This same proof establishes that T_n is a continuous map for the larger family of d.f.'s possessing the increasing failure rate average (IFRA) property (c.f., Barlow and Proschan (1975, p. 84)). However, we shall have no need for this greater generality; many of the results in this paper do not extend to this larger family.

Proportion 8. *For any set* $S \subseteq F$, $T_n(S^+) = T_n(S)^+$. *I.e.* T_n *maps the closure of* S *onto the closure of* $T_n(S)$.

Proof. Proportion 7 immediately implies $T_n(S^+) \subseteq T_n(S)^+$; the reverse inclusion depends on a standard argument based on tightness and relative compactness. □

Proposition 9 $T_n(G_n)^+ = T_n(G_n) + T_n(F_n)$, $T_n(H_n)^+ = T_n(H_n) + T_n(F_n)$. ($F_n$ *defined in Proposition* 1), *and* $T_n(F_n) = \sum_{i=1}^{n-1}\{T_n(G_i)+T_n(H_i)\}$ ($T_n(F_1)$ *is empty*) , $n \geq 1$.

Proof. The empty intersections indicated by set additions are all validated by using the latter part of Proposition 5 (together with the obvious fact that the families G_m and H_m, $m \geq 1$, are all disjoint). The proposition then follows immediately from Propositions 1 and 8. □

Proposition 10. *There exists a homeomorphic mapping (one-to-one and bicontinuous) from* $\mathbb{R}_+^n$ *onto* G_n, *and likewise for* H_n.

<u>Proof (sketch)</u> As noted in Section 2, the graph $L(G)$ of a d.f. $G \in G_n$ is a convex-shaped polygonal line which can be parameterized by n real parameters $y_1, \ldots, y_n$ representing the locations of nodes and the slopes of segments. Of these, $[\frac{n+1}{2}]$ are location parameters, and $[\frac{n}{2}]$ are slope parameters. If these parameters are assigned in an orderly way, moving from left to right, then the location parameters (say $y_1, y_3, \ldots$) form a strictly positive, strictly increasing sequence, and so do the slope parameters $(y_2, y_4, \ldots)$. A parameter $x = (x_1, \ldots, x_n)$ in $\mathbb{R}^n_+$ can be defined (not uniquely) by setting $x_1 = y_1$, $x_2 = y_2$, $x_3 = y_3 - y_1$, $x_4 = y_4 - y_2$, $\ldots$, $x_m = y_m - y_{m-2}$. It is intuitively clear that the map $x \mapsto G$ is a homeomorphism. For small changes in x result in small changes in $L(G)$, which result in small changes in G; and conversely. A similar argument applies to H_n. □

The following proposition is well-known among algebraic topologists. A proof is given by Greenberg (1967, pp. 81,82).

<u>Proposition 11 (invariance of domain)</u>. *Let* U *be an open connected subset of* $\mathbb{R}^n$ *and* $f: U \to \mathbb{R}^n$ *be a continuous and one-to-one function. Then* f(U) *is open and connected, and* f *is a homeomorphism onto* f(U) $(n \geq 1)$.

<u>Proposition 12</u>. *For each* $n \geq 1$, $T_n(G_n)$ *and* $T_n(H_n)$ *are simply connected open subsets of* $\mathbb{R}^n_+$, *and the mappings* $T_n: G_n \to T_n(G_n)$ *and* $T_n: H_n \to T_n(H_n)$ *are homeomorphisms. Thus* $T_n(G_n)$ *and* $T_n(H_n)$ *are open and connected subsets of* M_n, *necessarily in the interior of* M_n.

<u>Proof</u>. This follows immediately from Propositions 4, 7, 10 and 11 with $U = \mathbb{R}^n_+$. □

For $n \geq 2$, let M_n^* denote the set of points $c = (c_1,\ldots,c_n) \in \mathbb{R}_+^n$ for which there are d.f.'s $G \in G_{n-1}$ and $H \in H_{n-1}$ such that

$$(c_1,\ldots,c_{n-1}) = T_{n-1}(G) = T_{n-1}(H) \text{ and } T^{(r_n)}(G) < c_n < T^{(r_n)}(H). \qquad (5)$$

These sets are introduced so that an intricate topological argument can be isolated within the proof of the next proposition; it will eventually become clear that $M_n^* = M_n^\circ$. (See (8) below.)

<u>Proposition 13</u>. $M_n^* \subseteq T_n(G_n) \cap T_n(H_n)$, $n \geq 2$.

<u>Proof</u> Suppose $n \geq 2$ and $c = (c_1,\ldots,c_n) \in M_n^*$, so that (5) holds for some $G \in G_{n-1}$ and $H \in H_{n-1}$. It will be shown that $c \in T_n(G_n)$ by showing that there is a point $c^* = (c_1^*, \ldots, c_n^*) \in T_n(G_n)$ and a continuous path (parameterizable by a real variable) from c^* to c which does not intersect the boundary of $T_n(G_n)$. (That this implies that $c \in T_n(G_n)$ is easily verified.) The contemplated path consists of two straight line segments, the first from c^* to $(c_1^*,\ldots,c_{n-1}^*,c_n)$, and the second from the latter point to c. A similar argument can be employed to show that $c \in T_n(H_n)$.

Referring to (5), Proposition 7 and Proposition 12, one finds that there exists a small $\varepsilon > 0$ such that (i) an open ball B in $\mathbb{R}_+^{n-1}$, centered at $(c_1,\ldots,c_{n-1})$ of radius ε, is a subset of $T_{n-1}(G_{n-1}) \cap T_{n-1}(H_{n-1})$, and (ii) whenever $T_{n-1}(G') \in B$ for some $G' \in G_{n-1}$ and $T_{n-1}(H') \in B$ for some $H' \in H_{n-1}$, then $T^{(r_n)}(G') < c_n < T^{(r_n)}(H')$.

By Proposition 9, there exists a d.f. $G_0 \in G_n$ which is sufficiently close to G that $T_{n-1}(G_0) \in B$. Let $c^* = (c_1^*, \ldots, c_n^*) = T_n(G_0)$, and observe that

$(c_1^*,\ldots,c_{n-1}^*) \in B$. By (i), there exist d.f.'s $G^* \in \mathcal{G}_{n-1}$ and $H^* \in \mathcal{H}_{n-1}$ for which $T_{n-1}(G^*) = T_{n-1}(H^*) = (c_1^*,\ldots,c_{n-1}^*)$; and by (ii) ,

$$T^{(r_n)}(G^*) < c_n < T^{(r_n)}(H^*) \ . \tag{6}$$

Moreover, by Proposition 6,

$$T^{(r_n)}(G^*) < T^{(r_n)}(G_0)=c_n^* < T^{(r_n)}(H^*) \ . \tag{7}$$

Now suppose (for the purpose of obtaining a contradiction) that the line segment joining $(c_1^*, \ldots, c_{n-1}^*, c_n)$ and c^* intersects the boundary of $T_n(\mathcal{G}_n)$ at some point $(c_1^*, \ldots, c_{n-1}^*, x)$. This point is expressible as $T_n(F)$ for some $F \in \mathcal{F}_n$ since (the open set) $T_n(\mathcal{G}_n)$ contains none of its boundary points. (See Propositions 9 and 12.) However, x is some number between c_n and c_n^*; so by (6) and (7), $T^{(r_n)}(G^*) < T^{(r_n)}(F) = x < T^{(r_n)}(H^*)$ (strict inequalities). Consequently $F \neq G^*$ and $F \neq H^*$, even though $T_{n-1}(F) = (c_1^*, \ldots, c_{n-1}^*) = T_{n-1}(G^*) = T_{n-1}(H^*)$, and it follows from Proposition 6 that $F \notin \mathcal{G}_{n-1}^+ \cup \mathcal{H}_{n-1}^+ = \mathcal{F}_n$ (see Proposition 1), a contradiction. Thus the line segment joining $(c_1^*,\ldots,c_{n-1}^*, c_n)$ and c^* does *not* intersect the boundary of $T_n(\mathcal{G}_n)$.

Finally suppose (again for the purpose of obtaining a contradiction) that the line segment joining c and $(c_1^*, \ldots,c_{n-1}^*, c_n)$ intersects the boundary of $T_n(\mathcal{G}_n)$ at some point $(x_1,\ldots,x_{n-1}, c_n)$, (again)necessarily expressable as $T_n(F)$ for some $F \in \mathcal{F}_n$. It is easily seen that $(x_1,\ldots,x_{n-1}) \in B$, so that there exist d.f.'s $G_1 \in \mathcal{G}_{n-1}$ and $H_1 \in \mathcal{H}_{n-1}$ for which $T_{n-1}(G_1) = T_{n-1}(H_1) = (x_1,\ldots,x_{n-1})$. By (ii) above, $T^{(r_n)}(G_1) < c_n = T^{(r_n)}(F) < T^{(r_n)}(H_1)$, so that $F \neq G_1$ and $F \neq H_1$. Again, from Proportion 6, one obtains the contradiction $F \notin F_n$. Thus the line segment joining c and $(c_1^*,\ldots c_{n-1}^*, c_n)$ does *not* intersect the boundary of $T_n(G_n)$; and we have established that $c \in T_n(G_n)$, as required. □

Proposition 14 $T_n(F_n^c) \subseteq M_n^*$, $n \geq 2$.

Proof. (by induction) Suppose $c = (c_1,c_2) = T_2(F)$, where $F \in F_2^c$. Then $c_1 = T_1(G)$ when $G \in G_1$ is degenerate at c_1, and $c_1 = T(H)$ when $H \in H_1$ is exponential with mean $c_1/(\Gamma(r_1+1))^{1/r_1}$. By Proposition 6, $T^{(r_2)}(G) < T^{(r_2)}(F) < T^{(r_2)}(H)$. Thus, by definition, $c \in M_2^*$. Suppose now that the proposition holds for $n = m-1 \geq 2$, and that $c = (c_1,\ldots,c_m) = T_m(F)$, where $F \in F_m^c$. Since $F_m^c \subseteq F_{m-1}^c$, $T_{m-1}(F) \in M_{m-1}^* \subseteq T_{m-1}(G_{m-1}) \cap T_{m-1}(H_{m-1})$ (by the induction assumption and Proposition 13). Thus $(c_1,\ldots,c_{m-1}) = T_{m-1}(F) = T_{m-1}(G) = T_{m-1}(H)$ for some $G \in G_{m-1}$ and $H \in H_{m-1}$. Apply Proposition 6 again, as above, to show that $c \in M_n^*$. □

Proposition 15 M_n° *and* $T_n(F_n)$ *are disjoint, and thus* $M_n^\circ \subseteq T_n(F_n^c)$, $n \geq 1$.

Proof This is obvious for $n = 1$ since F_1 is empty. For fixed $n \geq 2$, suppose $T_n(F) \in M_n^\circ$ for some $F \in F$. Since M_n° is open (see Proposition 12), there is another d.f. $F' \in F$ such that $T_n(F') \in M_n^\circ$, $T_{n-1}(F') = T_{n-1}(F)$ and $T^{(r_n)}(F') < T^{(r_n)}(F)$. Proposition 6 rules out the possibility that $F \in G_{n-1}$ (the sense of the previous inequality is wrong); and, similarly, $F \in H_{n-1}$ is ruled out. Finally, Proposition 5 rules out the possibility that $F \in G_i$ or H_i for some i, $1 \le i \le n-2$, since $F' \neq F$ and for any such i, $T_{i+1}(F') = T_{i+1}(F)$. Thus $F \notin F_n$, and it follows that M_n° and $T_n(F_n)$ are disjoint. This implies $M_n^\circ \subseteq T_n(F_n^c)$ because $M_n^\circ \subseteq M_n = T_n(F) \subseteq T_n(F_n) \cup T_n(F_n^c) = T_n(F_n^c)$. □

The proofs of Theorems 1-5 will now be given.

Proof of Theorem 1. Associated with each point $T_n(F) \in M_n$, $F \in F$, there is a "ray" of points generated by $T_n(F_\theta)$, $0 < \theta < \infty$, emanating from the origin, where $F_\theta \in F$ is defined by $F_\theta(\cdot) = F(\cdot/\theta)$. Thus M_n is a cone. The fact that M_n is closed follows immediately from Proposition 8. □

Proof of Theorem 3. For $n \ge 2$, apply Propositions 15, 14, 13 and 12, in this order, to obtain

$$M_n^\circ \subseteq T_n(F_n^c) \subseteq M_n^* \subseteq T_n(G_n) \cap T_n(H_n) \subseteq T_n(G_n) \cup T_n(H_n) \subseteq M_n^\circ; \quad (8)$$

equality must hold throughout. For $n = 1$, one has

$$M_n = T_1(F) = T_1(F_1^c) = T_1(G_1) = T_1(H_1) \subseteq M_n^\circ \subseteq M_n \ . \quad □$$

Proof of Theorem 2. By Theorem 3, $\partial M_n = M_n - M_n^\circ = T_n(F) - T_n(F_n^c) \subseteq T_n(F_n)$; and by Proposition 15, $T_n(F_n) = T_n(F_n) \cap \partial M_n \subseteq \partial M_n$. Thus $\partial M_n = T_n(F_n)$. The remainder of the theorem is established in Proposition 9. □

Proof of Theorem 4. The first part of (1) follows immediately from Theorem 3 except for the uniqueness of G and H, which follows from Proposition 3. The remaining assertions follow directly from Proposition 6. □

Proof of Theorem 5 When $F \in \mathcal{F}_n^c$ the first assertion is essentially covered by Theorem 4; and when $F \in \mathcal{F}_n$, $F \in \mathcal{G}_n^+ \cap \mathcal{H}_n^+$. The inequalities described in (3) are similar to those described in Proposition 6. For instance, the first of these inequalities is a further consequence of the fact that "$\overline{F}$-$\overline{G}$ and k exhibit the same arrangements of signs". (See Proposition 2 and the proof of Proposition 6.) □

4. MOMENT INEQUALITIES

In this section we apply the theory developed in the previous section to obtain tight moment inequalities in a few specific situations.

Consider a r.v. X whose d.f. F is in the IFR family. We suppose that the moments EX^s, $s = r_1, \ldots, r_n$, are known, where $(r_1, \ldots r_n)$ is a fixed vector in $\mathbb{R}_+^n$ with $0 < r_1 < \cdots < r_n$, and we seek the best possible upper and lower bounds for the r-th moment EX^r for some $r > 0$. Theorem 5 informs us that this can be accomplished by first finding the unique d.f.'s $G \in \mathcal{G}_n^+$ and $H \in \mathcal{H}_n^+$ whose r_j-th moments, $j = 1, \ldots, n$, agree with those for F. Let Y and Z denote r.v.'s with the required d.f.'s G and H respectively. Then the inequalities

$$\min(EY^r, EZ^r) \le EX^r \le \max(EY^r, EZ^r)$$

provide the best possible bounds for EX^r, $r > 0$. In particular,

$$EY^r \le EX^r \le EZ^r \quad , \quad r > r_n, \tag{9}$$

or whenever there is an even number of $r_j \ge r$ $(j = 1, \ldots, n)$. The remarkable fact is that the distributions of Y and Z do not depend on the choice of r. Hereafter, we shall restrict our attention to the case $r > r_n$, so that (9) is applicable.

(a) *Finding G and H.* There are two cases to consider: (i) $F \in \mathcal{F}_n$ and (ii) $F \notin \mathcal{F}_n$. The first of these is easily disposed of. For when $F \in \mathcal{F}_n$, there is no other d.f. in $\mathcal{F}$ with the same r_j-th moments $j = 1, \ldots, n$. Thus $G = H = F$, and (9) reduces to equalities. F belongs to $\mathcal{F}_n$ if and only if its graph is a polygonal line, with a finite number of segments, which is determined by fewer than n parameters (one for each strictly positive finite slope, and one for each node away from the origin).

If $F \notin \mathcal{F}_n$, then, according to Theorem 4, $G \in \mathcal{G}_n$ and $H \in \mathcal{H}_n$. The graph of G, and hence G itself, is specified by n parameters which can be obtained as the unique solution of the n equations

$$EY^{r_j} = EX^{r_j} \quad , \quad j = 1, \ldots, n. \tag{10}$$

Likewise, H is specified by means of n parameters determined by

$$EZ^{r_j} = EX^{r_j} \quad , \qquad j=1, \ldots, n. \tag{11}$$

Solutions to (10) and (11) are guaranteed by Theorem 4. In a few situations the solutions can be obtained analytically. In any event, one can always obtain numerical solutions, and once the solutions are found then the bounds EY^r and EZ^r can be calculated.

(b) *Lower bounds when n = 1.* Here one is concerned with $\mathcal{G}_1$, the degenerate IFR distributions. Given EX^{r_1}, one can solve equations (10) and conclude that the r-th moment of the distribution degenerate at $\{EX^{r_1}\}^{1/r_1}$ is a lower bound for EX^r, $r > r_1$. It follows that $\{EX^r\}^{1/r}$ is nondecreasing in r, a fact already well-known to be true, in general, for nonnegative r.v.'s.

(c) *Upper bounds when n = 1* Here one is concerned with H_1, the exponential distributions. Given EX^{r_1}, equation (11) is satisfied when Z is exponentially distributed with mean $\{EX^{r_1}/\Gamma(r_1+1)\}^{1/r_1}$. Then (9) leads to the known conclusion (see Barlow and Proschan (1975, p. 120)) that $\{EX^r/\Gamma(r+1)\}^{1/r}$ is nonincreasing in $r > 0$, a fact which holds more generally for IFRA distributions (distributions which have an increasing failure rate average).

(d) *Upper bounds when n = 2 (new results).* Here one is concerned with H_2, the positively shifted exponential distributions; $H \in H_2$ takes the general form

$$H(x) = \begin{cases} 0 & 0 \leq x < c, \\ 1 - \exp\{-m(x-c)\} & x \geq c, \end{cases}$$

where m and c are (strictly) positive constants, and (11) can be expressed as

$$\int_0^\infty \left(\frac{u}{m} + c\right)^{r_j} e^{-u}\, du = EX^{r_j} \quad , \quad j = 1,2.$$

For the special calse $r_1 = 1$, $r_2 = 2$, these equations reduce to $c+m^{-1} = EX$ and $c^2 + 2cm^{-1} + 2m^{-2} = EX^2$, so that $c = \mu-\sigma$ and $m = \sigma^{-1}$, where $\mu = EX$ and $\sigma^2 = \text{var } X = EX^2 - E^2X$. Then (9) yields the tight upper bounds

$$EX^r \leq \int_0^\infty (\mu+\sigma(u-1))^r\, e^{-u}\, du \quad , \; r > 2 \; , \tag{12}$$

(This bound happens to be valid for *all* $F \in F$ whose mean and standard deviation are μ and σ, and not just for $F \notin F_2$. When $F \in F_2$, F is either degenerate ($\sigma = 0$, corresponding to $m = \infty$), or F is exponential ($\sigma=\mu$, corresponding to $c = 0$). In both cases, (12) holds as an equality for all $r > 2$.) In particular,

$$EX^3 \leq \mu^3 + 3\mu\sigma^2 + 2\sigma^3 \quad , \tag{13}$$

which can be expressed in the more appealing forms

$$\mu_3 \leq 2\mu_2^{3/2} \quad , \quad K_3 \leq 2K_2^{3/2} \quad , \tag{14}$$

where μ_j denotes the j-th central moment of F, and K_j (= μ_j) denotes the j-th cumulant of F, j = 1,2. Thus for IFR distributions, the "moment ratio" $\gamma_1 = \mu_3/\mu_2^{3/2}$ never exceeds two.

Since many of the known properties of IFR distributions are shared by the larger family of IFRA distributions, one might reasonably wonder whether the theory of moment spaces presented herein is extendable, in toto, to the family of IFRA distributions. No such extension is possible because there are IFRA distributions which have a moment ratio γ_1 larger than two. A simple example is the following: Let $F(x) = 1-e^{-x}$ for $0 \leq x < a$, = $1-e^{-2x}$ for $x \geq a$, which has an increasing failure rate average when $a > 0$. When a = .3, $\gamma_1 \doteq 2.17$.

For r = 4, (12) becomes

$$EX^4 \leq \mu^4 + 6\mu^2\sigma^2 + 8\mu\sigma^3 + 9\sigma^4 \quad . \tag{15}$$

More generally, if $r_1 = 1$ and $r_2 = s > 1$, then the analogue of (12) becomes

$$EX^r \leq \int_0^\infty (\mu + \alpha(u-1))^r e^{-u} du, \quad r > s, \tag{16}$$

where α is the unique solution of the equation

$$EX^s = \int_0^\infty (\mu + \alpha(u-1))^s e^{-u} du \ .$$

(e) *Lower bounds when $n = 2$.* Here one is concerned with G_2, the truncated exponential distributions; $G \in G_2$ takes the general form

$$G(x)= \begin{cases} 1 - e^{-mx} & 0 \le x < c, \\ 1 & x \ge c, \end{cases}$$

where m and c are (strictly) positive constants, and (10) can be expressed as

$$s\,\gamma(s,a) = m^s\, EX^s \quad , \quad s = r_1, r_2, \tag{17}$$

where $a = cm$ and $\gamma(s,a)$ denotes the incomplete gamma function defined by

$$\gamma(s,a) = \int_0^a u^{s-1}e^{-u}\,du \qquad , \qquad s > 0,\ a > 0.$$

When $r_1 = 1$ and $r_2 = 2$, (17) reduces to

$$1 - e^{-a} = mEX,\ 2 - 2(1+a)e^{-a} = m^2EX^2$$

and "a" can be obtained as the unique solution of

$$(1-e^{-a})^2\, EX^2 = \{2-2(1+a)e^{-a}\}E^2X\ . \tag{18}$$

Then, in terms of "a", (9) yields the tight lower bounds

$$EX^r \ge rm^{-r}\gamma(r,a) = r\mu^r(1-e^{-a})^{-r}\gamma(r,a) \qquad , \qquad r > 2,$$

where $\mu = EX$. In particular,

$$EX^3 \ge 6\mu^3(1-e^{-a})^{-3}(1-[1+a+a^2/2]e^{-a}). \tag{19}$$

and

$$EX^4 \ge 24\mu^4(1-e^{-a})^{-4}(1-[1+a+a^2/2+a^3/6]e^{-a})\quad .$$

(f) *A numerical example* Let X have a gamma distribution with shape parameter $\alpha = 1.5$ and scale parameter $\beta = 1$, which is an IFR distribution. By direct calculation,

$$EX = \frac{\alpha}{\beta} = 1.5, \quad EX^2 = \frac{\alpha(\alpha+1)}{\beta^2} = 3.75, \quad EX^3 = \frac{\alpha(\alpha+1)(\alpha+2)}{\beta^3} = 13.125 \quad .$$

Lower bounds for EX^3, based on a single moment, are given by

$$EX^3 \geq E^3X = 3.375 \quad \text{and } EX^3 \geq (EX^2)^{3/2} = 7.262 \quad ,$$

and upper bounds, based on a single moment, are given by

$$EX^3 \leq 6\ E^3X = 20.25 \quad \text{and } EX^3 \leq 6(EX^2/2)^{3/2} \doteq 15.405 \ .$$

An upper bound for EX^3, based on two moments (using (13)), yields $EX^3 \leq 13.8$, and a lower bound for EX^3, based on two moments (using (19) with $a = 2.36$ obtained from (18)), yields $EX^3 \geq 11.45$. By using two moments instead of (the best) one, the error in the lower bound drops from 45% to 13%, and the error in the upper bound drops from 17% to 5%. Finally, there is another upper bound for EX^3, available in the literature, which is based on EX and EX^2, namely $EX^3 \leq \frac{3}{2} E^2X^2/EX \doteq 14.06$; it has an error of about 7%. This inequality is based on the fact that $\lambda_r = EX^r/\Gamma(r+1)$, is logconcave in r when X has an IFR distribution. See Marshall and Olkin (1979).[1]

(g) *Some graphs of upper and lower bounds.* When two moments are held fixed, there is a range of possible values for a third moment. One way of describing this range is to fix the first of the two moments and to plot the range of the third as a function of the second. There is no loss of generality in letting

establishment of a particular scaling. For the four graphs below, the

[1]Marshall and Olkin (1979, p. 494) indicate that the logconcavity holds for $r \geq 1$. A careful check of the proof of this result reveals that it holds for all $r \geq 0$.

moment that is held fixed is always the first; three of the graphs plot the range of a third moment as a function of a second, and a fourth graph plots the range of a cumulant as a function of another cumulant. The mathematical basis for these graphs is given in subsections 4d and 4e above.

In Figure 5, the possible range of the third moment of an IFR distribution is plotted as a function of the second moment; the first moment is held fixed at unity. For the entire range of possible values of the second moment, from 1 to 2, the upper bound for the third moment never exceeds the corresponding lower bound for the third moment by more than 24%.

REGION OF SECOND AND THIRD MOMENTS
(FIRST MOMENT=1)

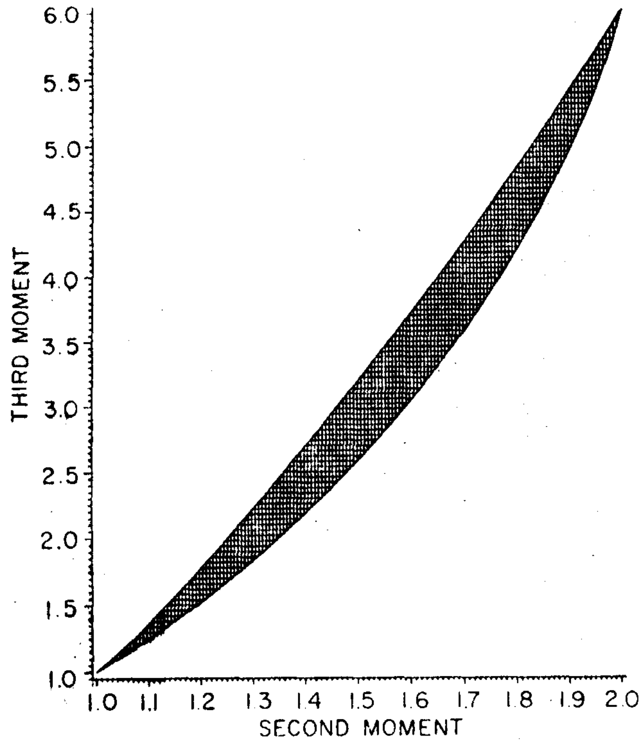

Figure 5

REGION OF SECOND AND FOURTH MOMENTS
(FIRST MOMENT=1)

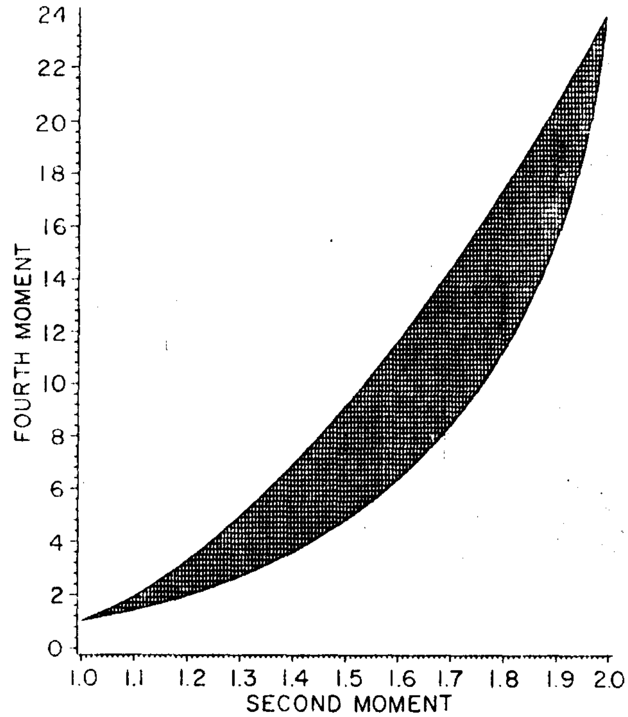

Figure 6

REGION OF THIRD AND FOURTH MOMENTS
(FIRST MOMENT=1)

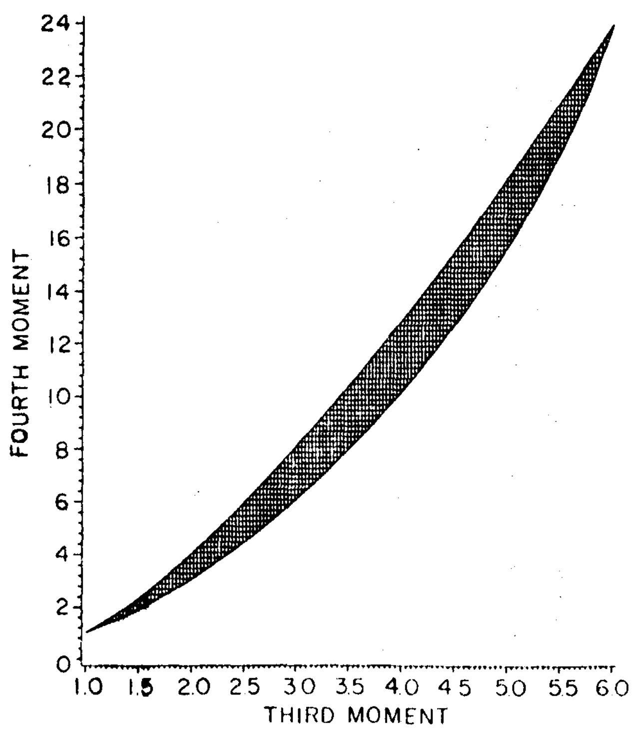

Figure 7

REGION OF SECOND AND THIRD CUMULANTS
(FIRST MOMENT=1)

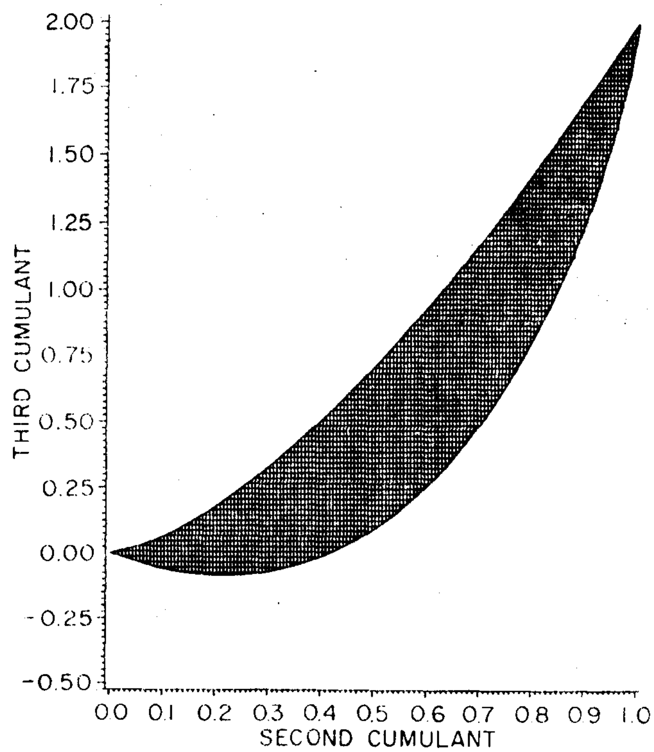

Figure 8

5. WEAK CONVERGENCE

In this section, we discuss the subject of weak convergence (convergence in law) within the IFR family of distributions. Since we found it convenient earlier to exclude from the family of IFR distributions the distribution which is degenerate at zero, a small price must be paid in the way we state the next theorem.

Theorem 6 . *Let* $\{F_m\}$ *be a sequence of* IFR *distribution functions.*

(i) *If* $F_m \to F$ *in law, then* F *is necessarily* IFR *or degenerate at zero, and*

$$\lim_{m\to\infty} \int_0^\infty x^r dF_m(x) = \int_0^\infty x^r dF(x) \tag{20}$$

for every $r > 0$.

(ii) *Conversely, if* (20) *holds for each integer* $r > 0$, *then* $F_m \to F$ *in law.*

Proof. Since $-\log\overline{F}_m(x)$ is convex on $I_{F_m} = \{x \ge 0: F_m(x) < 1\}$, it follows when $F_m \to F$ in law, that $-\log\overline{F}(x)$ is convex on $I_F = \{x \ge 0: F(x) < 1\}$. Consequently, F is IFR (when $I_F \neq \emptyset$) or degenerate at zero (when $I_F = \emptyset$). Proposition 7 proves (20) immediately when F is IFR, and the same argument can be used when F is degenerate at zero. This proves (i). The proof of (ii) requires a standard argument based upon tightness and relative compactness, together with the use of the fact that an IFR distribution is uniquely determined by its moment sequence. ∎

It is not necessary that (20) hold for *every* integer $r > 0$ in order that $F_m \to F$, but we do not know how much (ii) can be improved. Nevertheless, dramatic improvements are possible for certain types of F:

<u>Theorem 7.</u> *Let* $\{F_m\}$ *be a sequence of* IFR *distribution functions and* F *be any member of* $G_n \cup H_n$ *for some* $n \geq 1$. *If* (20) *holds for any* $n + 1$ *distinct positive values of* r, *then* $F_m \to F$ *in law.*

This result can be easily proved by using standard arguments involving tightness and relative compactness together with Proposition 5. It generalizes a theorem due to Obretenov (1977), which asserts that a sequence of IFR distributions, each having a unit mean, converges weakly to an exponential distribution if and only if the corresponding sequence of variances converges to 1. (His result corresponds to $F \in H_1$.)

It is apparent from Theorem 7, that weak convergence theory for the IFR family of distributions has quite peculiar features, not present in the general setting. The following theorem describes another unusual feature.

<u>Theorem 8</u> *Let* $\{F_m\}$ *be a sequence of* IFR *distribution functions.*

(i) *If* $F_m \to F$ *in the Kolmogorov metric, i.e., if* $\sup_{x\geq 0}|F_m(x)-F(x)| \to 0$ *as* $m \to \infty$, *then* $F_m \to F$ *in the strong sense of convergence in total variation:*

$$\sup\{|F_m(B) - F(B)|: B \textit{ Borel}\} \to 0 \textit{ as } m \to \infty . \tag{21}$$

(ii) *Thus if* $F_m \to F$ *in law and* F *is continuous, then* (21) *holds. If, in addition, each* F_m *is continuous, then each* F_m *has a p.d.f.* f_m, *and* F *has a p.d.f.* f *for which the convergence* $f_m \to f$ *occurs in the pointwise sense, and in the* L_1 *sense:*

$$\int_0^\infty |f_m(x) - f(x)|dx \to 0 \quad \textit{as } m \to \infty .$$

Proof. In order to prove (i), note that convergence in the Kolmogorov metric entails weak convergence, so that F is IFR or degenerate at zero. Now because of the convexity of $-\log\overline{F}_m$, the latter is absolutely continuous on $I_{F_m} = \{x:F_m(x)<1\}$, and hence has a nondecreasing density h_m on I_{F_m}. (See Roberts and Varberg ((1973), p. 10.) A little algebra then shows that the product $f_m = h_m\overline{F}_m$ is a p.d.f. of F_m on I_{F_m}. (If F_m is continuous, then f_m can be defined for all $x \geq 0$.) Likewise, F has a p.d.f. $f=h\overline{F}$ on $I_F = \{x \geq 0: F(x) < 1\}$. Now, according to Roberts and Varberg (1973, p. 20), $h_m \to h$ a.e. on I_F, and thus $f_m \to f$ a.e. on I_F. This is enough to establish (21) for all Borel sets $B \subseteq I_F$ (see, for instance, Billingsley (1968), p. 224). To establish (21) for all Borel sets B, special attention must be focused on the right endpoint x_0 of I_F, should such an endpoint exist. The establishment of (21) further requires that one show that $F_m(\{x_0\}) \to F(\{x_0\})$. This follows immediately from the fact that $F_m \to F$ in the Kolmogorov metric. To establish the first part of (ii), it need only be noted that when $F_m \to F$ in law and F is continuous, then the convergence occurs as well in the Kolmogorov metric. The remaining part of (ii) represent standard results when p.d.f.'s are defined and the convergence of $F_n \to F$ occurs in the total variational sense. (See Billingsley ((1968), p. 224.) □

To see that the mere convergence of $F_m \to F$ in law is not enough to establish (21), consider the simple example for which $I_F = [0,1)$, $F_m(x) = mx/2(m+1)$ for $0 \leq x < 1 + m^{-1}$ and $F_m(\{1+m^{-1}\}) = \frac{1}{2}$. It does not follow that $|F_m(\{1\}) - F(\{1\})| \to 0$.

6. ACKNOWLEDGEMENT

The authors gratefully acknowledge helpful conversations with various members of the mathematics faculty at the University of North Carolina whose knowledge of algebraic topology substantially exceeds that of the authors. In particular, thanks are due to James Damon, Sheldon Newhouse and James Stasheff.

REFERENCES

AHIEZER, N.I. and KREIN, M. (1962). *Some Questions in the Theory of Moments*. American Mathematical Society, Providence.

BARLOW, R.E. and MARSHALL, A.W. (1964a). Bounds for distributions with monotone hazard rate. *Ann. Math. Statist. 35*, 1234-1256.

BARLOW, R.E. and MARSHALL, A.W. (1964b). Bounds for distributions with monotone hazard rate II. *Ann. Math. Statist. 35*, 1258-1274.

BARLOW, R.E. and MARSHALL, A.W. (1967). Bounds on interval probabilities for restricted families of distributions. *Proc. Fifth Berkeley Symp. 3*, 229-257.

BARLOW, R.E. and PROSCHAN, F. (1975). *Statistical Theory of Reliability and Life Testing*. Holt, Rinehart and Winston, New York.

BARLOW, R.E., MARSHALL, A.W. and PROSCHAN, F. (1963). Properties of probability distributions with monotone hazard rate. *Ann. Math. Statist. 34*, 375-389.

BILLINGSLEY, P. (1968). *Convergence of Probability Measures*. Wiley, New York.

GREENBERG, M. (1967). *Lectures on Algebraic Topology*. Benjamin and Addison-Wesley, Reading

JOHNSON, N.L. and ROGERS, C.A. The moment problem for unimodal distributions. *Ann. Math. Statist. 22*, 433-439.

KARLIN, S. (1968). *Total Positivity I*. Stanford University Press, Stanford.

KARLIN, S., PROSCHAN, F. and BARLOW, R.E. (1961). Moment inequalities of Pólya frequency functions. *Pacific J. Math., 11*, 1023-1033.

KARLIN, S. and SHAPLEY, L.S. (1953). *Geometry of moment spaces*, Memoirs American Mathematical Society. No. 12.

KARLIN, S. and STUDDEN, W.J. (1966). *Tchebycheff Systems with Applications in Analysis and Statistics*. Interscience Publishers, Wiley, New York.

KREIN, M.G. (1959). *The ideas of P.L. Čebyšev and A. A. Markov in the theory of limiting values of integrals and their further development. A.M.S. Translations, 12*, 1-121.

REFERENCES
(cont'd)

MALLOWS, C.L. (1956). Generalizations of Tchebycheff's inequalities. *J. Roy. Statist. Soc. Series B 18*, 139-176.

MALLOWS, C.L. (1963). A generalization of the Chebyshev inequalities. *Proc. London Math. Soc. 13*, 385-412.

MARSHALL, A.W. and OLKIN, I. (1979). *Inequalities: Theory of Majorization and its Applications.* Academic Press, New York.

OBRETENOV, A. (1977). Convergence of IFR-distribution to the exponential one. *Compterendus de L'Académie, 30*, 1385-1387.

ROBERTS, A.W. and VARBERG, D.E. (1973). *Convex Functions.* Academic Press, New York and London.

SHOHAT, J.A. and TAMARKIN, J.D. (1943). *The Problem of Moments.* American Mathematical Society, Providence.