TRANSACTIONS

of the

EIGHTH PRAGUE CONFERENCE

on

INFORMATION THEORY, STATISTICAL DECISION FUNCTIONS, RANDOM PROCESSES

held at

Prague, from August 28 to September 1, 1978

VOLUME A

ACADEMIA

PUBLISHING HOUSE
OF THE
CZECHOSLOVAK ACADEMY OF SCIENCES

PRAGUE 1978

CZECHOSLOVAK ACADEMY OF SCIENCES

Scientific Editor: Academician Jaroslav Kožešník

TRANSACTIONS include contributions of authors reprinted directly in a photographic way.
For this reason the authors are fully responsible for the correctness of their text.

Printed in Czechoslovakia

POLYA-TYPE, SCHUR-CONCAVE AND RELATED PROBABILITY DISTRIBUTIONS

Rashid Ahmad, Abdulrahman M. Abouammoh

Glasgow

ABSTRACT

The objective of this paper is to give a unified and an extended structure of the above classes of distributions as well as their many applications. The behaviour of these families is investigated under most commonly occurring functional operations such as closure under convolution, a passage to a limit in the weak sense, reversal and mixing properties. The useful and smooth properties of unimodality, strong unimodality and their variants are found to hold for some subclasses of these applicable probability distributions.

1. INTRODUCTION AND PRELIMINARIES

The above concepts are fundamental and play a basic role in many probabilistic-statistical problems. In particular, these ideas occur in estimation, testing hypotheses and general statistical inference problems. These ideas have predominated in the past many investigations in statistical theory and applications, and solutions for considerably important problems have been sought through these concepts and their various forms, that is some stronger or weaker forms of these concepts. Such approach will be elucidated within the context of the paper. There is a large amount of work in these areas, for example Karlin (1956), Lehmann (1959), Birnbaum et al. (1966), Johnson and Kotz (1972), Barlow and Proschan (1975), Proschan and Sethuraman (1977) and many others.

In the sequel, a generalization of strong unimodality from one dimensional to n-dimensional space is suggested and it is shown that the class of n-dimensional strongly unimodal distributions is closed under convolution, passage to a limit weakly, mixing and reversal. Some other characterization properties are studied through a concavity structure of the underlying probability distributions.

For the classes M_s^n, $-\infty < s < n^{-1}$ of distributions which satisfy a certain convexity property, it turns out that for $s = 0$ the class M_s^n is equivalent to strongly unimodal class in an n-dimensional space. It is realized that the marginal measures (distributions) of an n-dimensional log-concave measures (distributions) is log-concave and hence Ibragimov 1956 result is implied as a special case in one-dimension of n-dimensional strongly unimodal distributions.

In connection with Polya type k (PT_k) distributions, it is observed that PT_2 is equivalent to the monotone likelihood ratio (MLR) property and that these distributions are strongly unimodal when the considered parameter in the definition is location parameter. Furthermore, it is found that PT functions have smooth behaviour as kernel of some integral transformations and are closed under convolution on the positive real line, mixing, reversal and passage to a limit weakly.

The class of increasing failure rate distributions (IFR) is considered and it is found that IFR and IFR average (IFRA) are closed under convolution, passage to a limit weakly but not under reversal and mixing. It is noticed that strongly unimodal PT_2, IFR and IFRA form an increasing sequence of classes of distributions. Thus, IFR is the smallest class containing all exponential distributions.

An n-dimensional density function which is Schur-Concave, that is it satisfies a permutation symmetric property is observed to exhibit essentially exchangeable or symmetric dependence structure. The class of Schur-Concave functions is closed under convolution. In addition we shall show some other structural properties of related exchangeable distributions or variables having some interesting applications. Finally, applications in the context of inferential aspects will be discussed.

2. LOG-CONCAVE PROBABILITY MEASURES AND DISTRIBUTIONS

The well known definition of unimodality as given by Khintchine in 1938 states that a real variable X has a <u>unimodal</u> distribution about a vertex a if its distribution function is convex on $(-\infty,a)$ and concave on (a,∞). An extensive treatment together with many other properties and results concerning these and related type of distributions for the univariate case is given by the authors (1977a). A slightly disturbing property of unimodal distributions was pointed out by Chung in 1953, that is the convolution of two unimodal distributions may not be unimodal. But if one considers the symmetric unimodal distributions, then such class is closed under convolution. Wintner in 1938, Ibragimov in 1956 adapted the strongly unimodal concept in order to preserve the unimodality property under convolution. A distribution F is <u>strongly unimodal</u> if its convolution with any unimodal distribution is again unimodal. Ibragimov gave an interesting characterization of such distributions, namely, a distribution F(x) is strongly unimodal if and only if log f(x) is concave where f(x) is the density of F(x). Such a density is absolutely continuous within the range of its definition and $[\log f(x)]' = f'(x)/f(x)$ is nonincreasing function in x. It can be shown that the normal, exponential, Wishart and uniform distributions are strongly unimodal. Moreover, it follows that if f(x) is symmetric density which is not log-concave then there exists a unimodal distribution with density g, say, such that the convolution f * g is not unimodal. There is another characterization result of strongly unimodal distributions, that is a positive twice differentiable density f is strongly unimodal if and only if

(2.1) $f'^2 > f f''$

and this relation (2.1) can be easily derived from the relation

$(\log f)'' = [(f''f-f'^2)/f^2] \leq 0.$

Next, we shall show the closure of strongly unimodal class of distributions under mixing. Let f and g be two strongly unimodal densities and $\alpha \in (0,1)$, then by inequality (2.1) one has

$$(2.2) \quad (\alpha f'+(1-\alpha)g')^2 - (\alpha f''+(1-\alpha)g'')(\alpha f+(1-\alpha)g) \geq$$
$$\geq \alpha(1-\alpha)(2f'g'-fg''-gf'')$$
$$\geq fg(f'/f-g'/g)^2 \geq 0.$$

It can be easily realized that such class is closed under convergence to a limit weakly. In fact, Lapin has shown that the class of unimodal distributions is closed under the weak convergence, see Lukacs (1970, p. 97). Furthermore, if we define F(-x) to be the reversal of F(x) then it is not difficult to see that all distributions which are unimodal or strongly unimodal are closed under reversal property. Before we give the multivariate structure we summerize the above results in the following.

Theorem 2.1. The class of one-dimensional strongly unimodal distributions is closed under convolution, mixing, passage to a limit in the weak sense and reversal.

Now, we give the main results of log-concave families of probability measures in R^n. A probability measure P on R^n is log-concave if for all open convex sets $A, B \subset R^n$ and $\alpha \in [0,1]$, we have

$$(2.3) \quad P(\alpha A+(1-\alpha)B) \geq (P(A))^\alpha \, (P(B))^{1-\alpha}.$$

It is clear that (2.3) is also true for all closed convex sets A, B, since any closed convex set is the limit of a decreasing sequence of open convex sets. Another consequence of (2.3) is the inequality

$$P\{(\alpha X+(1-\alpha)Y) \in A \cup B)\} \geq (P(X \in A))^\alpha \, (P(Y \in B))^{1-\alpha}$$

for some vectors $\underline{x}$ and $\underline{y}$ in R^n.

Borell (1975), Prékopa (1973) and Rinott (1976) have dealt with the necessary and sufficient conditions for a density function f(x) in R^n which generates [illegible]-bability measure P that satisfies a convexity property of the type

$$(2.4) \quad P(\alpha A+(1-\alpha)B) \geq \begin{cases} [(\alpha P(A))^s+((1-\alpha)P(B))^s]^{1/s} & \text{if } s \in (-\infty,0) \text{ or } = \frac{1}{n} \\ \min(P(A),P(B)) & \text{if } s = -\infty \\ (P(A))^\alpha \, (P(B))^{1-\alpha} & \text{if } s = 0. \end{cases}$$

In case s = 1/n, (2.4) holds for Lebesgue measure and is known as the Brunn-Minkowski inequality. A probability measure P in R^n on its corresponding distribution is said to belong to class M_s^n if it satisfies relation (2.3). The class M_s^n is in a sense a generalization of the class of log-concave measures. It is noted that for $s_1 \geq s_2 \geq \ldots \geq -\infty$, one has $M^n_{-\infty}$ is the largest class and $M^n_{s_1} \subseteq M^n_{s_2} \subseteq \ldots \subseteq M^n_{-\infty}$. It is also clear that $M^n_s = \phi$ if $s > 1/n$. Now we shall give the following characterization result which has been proved in a special form by Rinott (1976).

Theorem 2.2. Let P be a probability measure generated by an n-dimensional density f, that is $P(B) = \int_B f(x)dx$, $x \in R^n$, for any Borel set $B \subset R^n$. Then $P \in M^n_s$ if and only if there exist a version h of the density f such that $h^{s/(1-sn)}$ is convex if $s \in [-\infty, 0)$, log h is concave if $s = 0$ and $h^{s/(1-sn)}$ is concave if $s \in (0, 1/n)$.

Proof: The necessary part is easily obtainable, see for example the geometrical argument in Rinott (1976). To prove the sufficiency part, let $s = 0$ and the probability measure P generated by some density f be log-concave. Let $S(r,x)$ to be the sphere in R^n with radius $1/r$ and centre $x \in R^n$. Define $f(r,x) = |S^{-1}(r,x)| \int f(y)dy$, $y \in R^n$, and the integral is taken over $S(r,x)$ and $|S(r,x)|$ is Lebesgue measure. Hence, the log-concavity of P implies that $f(r,\alpha x+(1-\alpha)y) \geq (f(r,x))^{\alpha}(f(r,y))^{1-\alpha}$. Take $h = \liminf_{r\to\infty} f(r,x)$. Thus, h is log-concave and f = h almost everywhere.

Let us now call the density function f defined on R^n to be multivariate strongly unimodal if it is log-concave. In what follows, we shall investigate the behaviour of the class of multivariate strongly unimodal distributions under some functional operations. Denote such a class by U^n where n refers to the underlying dimension. For the proof of the theorem below we need the following lemmas.

Lemma 2.1. (Prékopa (1973, p. 337)). Let $P_1, P_2, \ldots, P_k$ be k probability measures on R^n and let $d\mu(z) = \sup_{\alpha_1 x_1 + \ldots + _k x_k = z} dP_1(x_1) \ldots dP_k(x_k)$, $z \in R^n$, where $\alpha_i > 0$ are constants and $\alpha_1 + \ldots + \alpha_k = 1$. Then $\mu(z)$ is a probability measure and furthermore

$$\int_{R^n} d\mu(z) \geq [\int_{R^n} (dP_1(x_1))^{1/\alpha_1}]^{\alpha_1} \ldots [\int_{R^n} (dP_k(x_k))^{1/\alpha_k}]^{\alpha_k}.$$

Lemma 2.2. (Parthasarathy (1967 p. 40)). Let $\{P_k\}$ be a sequence of probability measures defined on a metric space (Ω, m). Then

(i) $\limsup_{k\to\infty} P_k(A) \leq P(A)$ for every closed set $A \subset \Omega$

(ii) $\lim_{k\to\infty} P_k(B) = P(B)$ for every Borel set B whose boundary has P-measure zero.

Theorem 2.3. The class U^n of distributions is closed under convolution, mixing, reversal and passage to a limit weakly.

Proof: We proceed the proof of various parts in the same order as in the statement of the theorem. Let P_1, P_2 be in U^n, we need to show $\int_{R^n} P_1(x-y)d\,P_2(y)$ is in U^n, that is it is log-concave for x, $y \in R^n$. Since $P_1(x-y)d\,P_2(y)$ is log-concave in R^{2n}. Then by taking $y = (y_1, y_2)$ and $x = (x_1, x_2)$ and applying lemma 2.1. respectively, the result follows.

The proof of the mixture (convex combination) of P_1 and P_2 being in U^n can be obtained by (2.3) and an argument similar to (2.2) which can be applied by taking the partial derivative with respect to the components of the underlying vectors, and hence U^n is closed under mixing. Furthermore, it is easily seen from the definition of U^n that it is closed under reversal. To prove the last part of theorem that U^n is closed under passage to a limit weakly, let $\{P_k\}$ be a sequence of probability measures in U^n for every $k \geq 1$ and P_k converges weakly to a probability measure P. Now, we want to show that P is in U^n, that is, it satisfies (2.3). But (2.3) is satisfied whenever A, B are closed convex non-empty sets whose boundaries have P-measure zero and hence by using lemma 2.2., $P(A) = \lim_{k\to\infty} P_k(A)$ and $P(B) = \lim_{k\to\infty} P_k(B)$. Since any open convex set is the limit of an increasing sequence of closed convex non-empty sets whose boundaries have P-measure zero. Therefore, $P \in U^n$ for any open convex sets A and B. This proves the theorem.

Corollary 2.1. If a distribution $F(x) \in U^n$, $x \in R^n$, then any marginal distribution $F(x_1,\ldots, x_k) \in U^k$ where $k = 1,\ldots, n$.

Remark: the multivariate unimodality has been defined and discussed in many different ways such as generalized unimodal, linear unimodal, monotone unimodal, central convex unimodal and the last two concepts are included in multivariate symmetric unimodality, see Kanter (1977), Ahmad and Abouammoh (1977a) and references mentioned therein for various characterization results.

3. STRUCTURE OF PT AND MLR DISTRIBUTIONS

The smoothness and the nice beahviour of PT and MLR distributions is discussed to examplify the structure and the applications such as in decision theory and other inferential problems in statistics.

The family of distributions $F(x,\theta)$ (or their densities $f(x,\theta)$) of real random variables X depending on a real parameter θ is said to belong to the class PT_k if

$$(3.1) \qquad \Delta_k = \begin{pmatrix} f(x_1,\theta_1) & \cdots & f(x_1,\theta_k) \\ \vdots & & \\ f(x_k,\theta_1) & \cdots & f(x_k,\theta_k) \end{pmatrix}$$

is semi-positive definite matrix for every $k \geq 1$ and all $x_1 < x_2 <\ldots< x_k$ and $\theta_1 < \theta_2 <\ldots< \theta_k$. The family $F(x,\theta)$ belongs strictly to PT_k if Δ_k in (3.1) is positive definite matrix. If $F(x,\theta)$ belongs to PT_k for every $k = 1,2,\ldots$, then $F(x,\theta)$ belongs to PT_∞. The distribution $F(x,y)$ of the two real variables ranging over linearly ordered one-dimensional sets X and Y respectively is said to be totally positive of order k (TP_k) if $\Delta_k(x,y)$ is semi-positive definite, that is, $\Delta_i(x,y) \geq 0$ for all i, $1 \leq i \leq k$, and it is called strictly TP_k if $\Delta_i(x,y) > 0$ for all i, $1 \leq i \leq k$. In fact the PT_k distributions give two familiar and interesting cases for $k = 1,2$.

These are, the family $F(x,\theta)$ is PT_1 that is if and only if $f(x,\theta) \geq 0$ for every x and θ, therefore every distribution identified by a parameter is PT_1 and the family $F(x,\theta)$ is PT_2 if and only if

(3.2) $\quad f(x_1,\theta_1)f(x_2,\theta_2) - f(x_1,\theta_2)f(x_2,\theta_1) \geq 0.$

The later case arises in many inference problems in applied statistics, and any family of distributions whose densities satisfy (3.2) are said to have MLR. One may notice from the definition of PT_k classes of distributions that $PT_\infty \subset \ldots \subset PT_2 \subset PT_1$ and it can be also shown that the exponential family, Lehmann (1959, p. 115), non-central F, noncentral t and noncentral chi-square, Karlin ((1956), and some other families of distributions belong strictly to PT_∞ and hence they have MLR property. In other words, most of the distributions used in statistical inference are PT. However, the most notable example of a distribution which is not PT is Cauchy with density $f(x,\theta) = \{\pi[1+(x-\theta)^2]\}^{-1}$.

It was found that PT functions have very nice property when they are used as kernel of transformation, for example, see Karlin (1957). If $f_1(x,\theta)$ belongs to PT_∞ and is nth differentiable with respect to x for all θ, F_3 is some distribution associated with finite measure, $f_2(\theta)$ is a function of θ which has n sign changes and $f_2(x) = \int f_1(x,\theta)\, f_2(\theta) d\, F_3(\theta)$ is nth differentiable with respect to x inside the integral, then $f_2(x)$ has at most n sign changes. Therefore, if f_1 and f_2 are two continuous differentiable densities of two independent random variables X and Y respectively such that $f_1(x-\theta)$ is strictly PT_∞ and f_2 has m modes, then the density of $Z = X+Y$ that is $f_3(z) = \int f_1(t)\, f_2(z-t)dt$ has at most m modes. Hence, the concavity (or convexity) property of a function is preserved under convolution with any strictly PT_∞ distribution. Furthermore, if f_1^* is the nth convolution of f_1, where f_1 is PT_∞ and f_2 is concave (convex) then $g(.,n) = \int f_1^{*n}(x)\, f_2(x)dx$ is concave (convex). Now we shall give the following lemma which can be proved by the basic composition formula and the direct product of matrices.

Lemma 3.1. The class of PT_k, $k = 1,2,\ldots$ distributions of non-negative random variables is closed under convolutions.

Theorem 3.1. The class of PT_k; $k = 1,2,\ldots$ distributions is closed under, mixing and convergence to a limit in the weak sense.

Proof: It is clear that PT class is closed under reversal since Δ_k defined by (3.1) is semi-definite for any $x_1 < x_2 < \ldots < x_k$ whether positive or negative values of x's. Now, let f_1 and f_2 belong to PT_k, $k \geq 1$ and $f = \alpha\, f_1+(1-\alpha)\, f_2$, $0 \leq \alpha \leq 1$. Therefore, by writing Δ_k in the convex combination form, also by taking it as a summation of two matrices the first of f_1 and the second of f_2, and since $\alpha^k \geq 0$, $(1-\alpha)^k \geq 0$ one has PT_k for $k \geq 1$ is closed under mixing. Finally, PT_k, $k \geq 1$, class is closed under convergence to a limit weakly is implied from the fact that the limit of a matrix is defined by the limit of each element of such matrix, which themselves are of PT form. This completes the proof.

As a consequence from the above we give the following.

Corollary 3.1. The class of distributions with MLR propetty is closed under reversal, mixing and convergence to a limit weakly.

Next, we give two weaker forms of MLR property which are defined by relation (3.2). These are (i) the parameter $\theta \in \Theta \subset R$ is merely a location parameter that is for $x_1 < x_2$ and $\theta_1 < \theta_2$ one has

(3.3) $f(x_1-\theta_1)\, f(x_2-\theta_2) - f(x_1-\theta_2)\, f(x_2-\theta_1) \geq 0,$

and (ii) if $x_1, x_2, x_3, x_4 \in R$ or integers and $\theta_1, \theta_2 \in \Theta$ such that $x_3 < x_1 < x_4$ and $\theta_1 < \theta_2$ we have

(3.4) $f(x_1,\theta_1)\, f(x_2,\theta_2) - f(x_4,\theta_1)\, f(x_3,\theta_2) \geq 0.$

Then one can see that (3.2) implies that $f(x_1,\theta_2)/f(x_1,\theta_1)$ is nondecreasing in x and $f(x_2,\theta_1)/f(x_1,\theta_1)$ is nondecreasing in θ_1 that is $f(x+h,\theta_1)/f(x,\theta_1)$ is nondecreasing in θ_1 for all x and $h > 0$. Also, (3.3) implies $f(x+h,\theta_1)/f(x,\theta_1)$ is nonincreasing in x; $h > 0$ and it implies that f is log-concave function. Further, (3.4) is equivalent to (3.2) if $x_1 = x_3$ and $x_2 = x_4$, and any class of functions satisfying (3.4) is closed under convolution see Ghurye and Wallace (1959). Now we summarize the closure of MLR class under convolutions as below.

Theorem 3.2. Let f_1 and f_2 be two density functions and $f = f_1 * f_2$. Then (3.4) is satisfied by f if it is satisfied by f_1 and f_2; (3.3) is satisfied by f if it is satisfied by f_1 and f_2 and (3.2) is satisfied by f if it is satisfied by f_1 and f_2 for non-negative random variables.

4. IFR TYPE DISTRIBUTIONS

It is assumed sometimes that the distribution of the future life (life distribution) of a device remains the same regardless of the time while it was in use, which is usually known in statistical term by 'new is the same as used', and is characterized by a life distribution of lack of memory such as exponentials. Such a distribution is said to represent no wear phenomenon. Similarly, other classes of distributions may represent the wear out phenomenon, that is new better than used (NBU) and the durability phenomenon that is new worse than used (NWU). Many authors have tackled the problem of finding classes of distributions which reflect these phenomena. To answer this problem Birnbaum et al. (1966) introduced the class of distributions with IFR which was also discussed by Barlow and Proschan (1975), A-Hameed and Proschan (1973), Black and Savits (1976) and many others.

Let the survival probability or reliability be $\bar{F}(x) = P(X > x)) = 1-F(x)$ which is the complement of the life distribution $F(x)$. The conditional reliability for the remaining life given that the device has survived to age t is $\bar{F}(x|t) = \bar{F}(x+t)/\bar{F}(x)$ if $\bar{F}(t) > 0$ and 0 if $\bar{F}(t) = 0$. Similarly the conditional probability failure during a time x for a device of age t is $F(x|t) = (F(x+t) - F(t))/\bar{F}(t) = 1-\bar{F}(x|t)$ if $\bar{F}(t) > 0$ and 0 if $\bar{F}(t) = 0$. It is noticed that no wear characteristic means unfailed device is treated as new that is for all x, $t > 0$, $\bar{F}(x|t) = \bar{F}(x|0)$, that is the class which satisfies the functional form $\bar{F}(x+t) = \bar{F}(x)\,\bar{F}(t)$ - in other words the class of exponential survival distributions $F(x) = \exp(-\lambda x), \lambda \in [0,\infty]$.

Now we may obtain the conditional failure rate r(t) at time t by

$$r(t) = \lim_{x\to 0} (1/x)\ \{[F(x+t)-F(x)]/\bar{F}(t)\} = f(t)/\bar{F}(t)$$

where $\bar{F}(t) > 0$ and f(t) is the density function of F(t). The failure function (or the cumulative failure rate) is

$$R(x) = \int_0^x r(t)dt = -\log \bar{F}(x)$$

and hence $\bar{F}(x) = \exp{-R(x)}$. A distribution function F is <u>IFR</u> if $F(x|t)$ is decreasing in t where $t > 0$ and $x > 0$ and this means r(t) is increasing function in t. A distribution function is said to be a decreasing failure rate (<u>DFR</u>) if $F(x|t)$ is increasing in t for all t and $x > 0$ and this implies that r(t) is decreasing. The inverse of the failure rate function $r^{-1}(t)$ is known by Mills' ratio and has been studied and tabulated by many authors, see Johnson and Kotz (1972) for references. Some wider classes than IFR and DFR are the classes of distributions with IFRA and DFRA. A distribution F belongs to <u>IFRA (DFRA)</u> class if R(x)/x is increasing (decreasing) and hence $[\bar{F}(x)]^{-1/x}$ is increasing (decreasing). Therefore, F is IFRA if and only if $-\log \bar{F}$ is star-shaped, where the non-negative function g on $[0,\infty)$ with $g(0) = 0$ is star-shaped if $x^{-1}g(x)$ is increasing in $x \in (0,\infty)$ or equivalently $g(\alpha x) \leqslant \alpha\, g(x)$ for $0 \leqslant \alpha \leqslant 1$, $0 \leqslant x \leqslant \infty$. Thus IFRA (DFRA) is characterized by $\bar{F}(\alpha x) \geqslant (\leqslant)\ \bar{F}^{\alpha}(x)$ $0 \leqslant \alpha \leqslant 1$, $x \geqslant 0$.

Next, we give the main classification of life distributions. A distribution F belongs to NBU if $\bar{F}(x) > \bar{F}(x+t)/\bar{F}(t)$ for all x, $t \geqslant 0$, that is $-\log \bar{F}$ is super-additive where the non-negative g defined on $[0,\infty)$ with $g(0) = 0$ is superadditive if $g(x+y) \geqslant g(x)+g(y)$ for x, $y \geqslant 0$. Similarly F belongs to NWU if $F(x) \leqslant F(x+t)/F(t)$ that is $-\log F$ is subadditive where a non-negative function g defined on $[0,\infty)$ with $g(0) = 0$ is subadditive if $g(x+y) \leqslant g(x)+g(y)$ for $x,y \geqslant 0$. The distribution F is said to belong to the class of new better (worse) than used in expectation <u>NBUE</u> (<u>NWUE</u>) if $\int_0^\infty (F(x+t)/\bar{F}(t))dx \leqslant (\geqslant) \int_0^\infty \bar{F}(x)dx < \infty\ (> -\infty)$ for all $t > 0$. The distribution F is said to belong to the class of decreasing (increasing) mean residual life <u>DMRL (IMRL)</u> if $\int_0^\infty (\bar{F}(x+t)/\bar{F}(t))dx$ is decreasing (increasing in t), that is, the residual life of an unfailed device of age t has mean which is decreasing (increasing) in t. Finally, if we denote the class of functions which satisfy relation (3.3) by T_2, that is PT_2 with extra condition $f(t+x)/f(t)$ is decreasing in t one can establish the result below.

<u>Theorem 4.1.</u> For the above classes one has: (i) $T_2 \subset$ IFR $\subset$ IFRA $\subset$ NBU $\subset$ NBUE, (ii) IFR $\subset$ DMRL $\subset$ NBUE, (iii) relation (i) and (ii) are the only existing among these classes.

The IFR class is closed under convolution, see Barlow and Proschan (1975, p. 100) and recently, Black and Savits (1976) proved that IFRA class is closed under convolution so one may ask whether other classes of theorem 4.1. are closed under convolution. However, it is realized that IFR is not closed under mixing, and a simple example for this is to note that mixture of two IFR exponential may not be IFR. Moreover, none of these classes is closed under reversal. It is also noticed that DFR class if not closed under convolution in general and this is proved by the counter example; let $F_1 = F_2 = G_\alpha$ where g_α is Γ-density with shape parameter $\alpha \varepsilon (\frac{1}{2},1)$, i.e. $g_\alpha(t) = \Gamma^{-1}(\alpha)\lambda^\alpha t^{\alpha-1}e^{-\lambda t}$ $t > 0$, then F_1 and F_2 are DFR distributions but $F_1 * F_2 = G_{2\alpha}$ which is Γ-distribution with shape parameter more than one and is not DFR distribution. However all classes of distributions included in this section are closed under convergence to a limit weakly.

5. EXCHANGEABILITY AND MAJORIZATION IN DISTRIBUTIONS

Here, we shall explore many results concerning the concepts of exchangeability, rearrangement, majorization and Schur concavity (or convexity) along with many other related ideas arising inevitably in probabilistic and statistical models. We shall discuss the basic properties of these concepts and establish some results concerning preservation property under some functional transforms in a similar manner to those in the earlier sections.

The random variables $X_1,\ldots, X_k$ are called exchangeable if k! permutations $(X_{t_i},\ldots, X_{i_k})$ have the same k-dimensional distribution. The sequence of random variables $X_1,X_2,\ldots$ is said to be spherical exchangeable if there exists a function g on the positive real line such that for each finite set $(i_1,\ldots, i_k)$ of natural numbers the joint characteristic function ϕ of $X_{i_1},\ldots, X_{i_k}$ satisfies $\phi(t_1,\ldots, t_k) = E\exp(i\Sigma t_jX_j) = g(\sum_{j=1}^{k} t_j^2)$. Clearly, each spherical exchangeable process is exchangeable. However, the exchangeable random variables $X_1,\ldots, X_k$ with probabilities of the form $P(X_1-\theta_1,\ldots, X_k-\theta_k) = P(X \varepsilon A+\theta)$, θ is parameter vector, often exhibit a monotonicity property in values of θ partially ordered according to majorizations. Notice that, we denote by X,Y,... and $X_1,X_2,\ldots, Y_1,Y_2,\ldots$ as the random vectors and random variables or components respectively. In fact, this concept has been studied by the authors (1977b) in the context of constructing A-infinitely divisible classes where A refers to the symmetric dependence of the variables.

It was established by Hardy et al. (1952, p. 49) that an n-dimensional vector X is said to be majorized by the vector Y if by rearrangement of the components to obtain $x_1 \geq x_2 \geq\ldots\geq x_n$, $y_1 \geq y_2 \geq\ldots\geq y_n$ one has

$$(5.1) \quad \sum_{j=1}^{k} x_j \leq \sum_{j=1}^{k} y_j,\ k = 1,2,\ldots, n-1 \text{ and } \sum_{j=1}^{n} x_j = \sum_{j=1}^{n} y_j,$$

and we denote this by $X \overset{*}{<} Y$ if relation (5.1) is satisfied. A function f for which $X \overset{*}{<} Y$ implies $f(x) \geq (\leq) f(y)$ is called <u>Schur-concave (convex)</u> and such functions are permutation symmetric, that is, invariant under permutations of the components of the underlying vectors. Therefore, f(x) is Schur-concave implies that the random variables $X_1, \ldots, X_n$ are exchangeable. Thus a differentiable function f(x) of exchangeable random variables is Schur-concave (convex) if and only if

(5.2) $$\left(\frac{\partial f(x)}{\partial x_i} - \frac{\partial f(x)}{\partial x_j}\right)(x_i - x_j) \leq (\geq)\ 0 \text{ for all } i \neq j,$$

see Schur (1923). The case $X \overset{*}{<} Y$ can be expressed by $X = DY$ for some doubly stochastic matrix D. It can be realized that for any vector x, $(\sum_{i=1}^{n} x_i/n)(1,1,\ldots,1) \overset{*}{<} (x_1, x_2, \ldots, x_n)$ and therefore, whenever $\sum_{i=1}^{n} x_i$ is fixed, Schur-concave function attains a maximum (modal) point when the components are equal. Now we give the following lemma due to Marshall and Olkin (1974).

<u>Lemma 5.1.</u> Let f(x) be a Schur-concave function and consider a Lebesgue-measurable set $A \subset R^n$ such that

(5.3) $y \in A$ and $x \overset{*}{<} y$ implies $x \in A$.

Then $P(X \in A+\theta) = \int_{A+\theta} f(x)dx$ is Schur-concave function of θ, where θ is some parameter vector.

In fact, condition (5.3) can be satisfied for every convex set A of exchangeable random variables since $x \overset{*}{<} y$ implies $x = Dy$ for some doubly stochastic matrix D and the set of doubly stochastic matrices is a convex hull of the permutation matrices, whereas (5.3) implies neither convexity nor measurability of A. Moreover, if two sets satisfy (5.3), then so does their union. Actually, Mudholkar (1966) established lemma 5.1 where he generalizes a result of Anderson (1955), see Kanter (1977), but with additional requirement on the set A that is A and $\{y: f(y) \geq c\}$ are convex for each constant c. Also $\{y: f(y) \geq c\}$ is convex, i.e. unimodal in Anderson's sense and f(y) is exchangeable implies condition (5.3). For some related results on exchangeability see Hewitt-Savage (1955) and Ahmad (1974, 1975). Now we give the following interesting result.

<u>Theorem 5.2.</u> The class of Schur-concave (convex) functions is closed under reversal, passage to a limit weakly, mixing and convolution.

<u>Proof</u>: We shall give the proof for Schur-concave functions, whereas a similar argument can be carried out for the Schur-convex case. Clearly, the reversal property is valid that is if f(x) is Schur-concave, then so is f(-x). Now let $\{f_k\}$ be a sequence of Schur-concave functions. Then we can have for any set A satisfying lemma 5.1, that $|f_k| \leq h$ for each k where h is integrable function on A+θ. Next, let f_k converge weakly to a function f then by Lebesgue Dominated Convergence Theorem and lemma (5.1) one gets f as Schur-concave. The closure under mixing is shown if one realizes that the mixture is Schur-concave function of θ.

Finally to show that the class of Schur-concave functions is closed under convolutions, let f_1 and f_2 be two Schur-concave functions, then $f_2(-x)$ is also Schur-concave and we need to prove that

$$(5.4) \qquad f(.,\theta) = \int_{R^n} f_1(x)\, f_2(x-\theta)dx$$

for some parameter θ, is Schur-concave in θ. But by using lemma 5.1,

$\int_{A+\theta} f_2(-x)dx = \int_{R^n} I_A(-x)\, f_2(\theta-x)dx$ is Schur-concave in θ. Now approximate $f_1(x)$ by an increasing sequence of simple functions $h_k = \Sigma\, \alpha_i\, I_{A_i}$ where $\Sigma\, \alpha_i = 1$ and the sets A_i satisfy lemma 5.1. Hence by using Lebesgue Monotone Convergence Theorem the required result follows.

One may also see that if f(x) is an exchangeable Schur-concave density function and h is non-negative, exchangeable and Schur-function, then $Eh(X-\theta)$ and $P\{h(X-\theta) \geq c\}$ are Schur-function in θ. The following lemma is due to Proschan and Sethuraman (1977), where they show the closure property of Schur-functions under a certain integral transformation.

<u>Lemma 5.2.</u> Let $f_1(x)$ be Schur-concave (convex) function and $f_2(x,\theta)$ is TP_2 and satisfy the semigroup property i.e. for $\theta_1,\theta_2 > 0$: $f_2(x,\theta_1+\theta_2) =$

$\int f_2(x,\theta_1)\, f_2(x,\theta_2-y)dy$, where $0 < \theta, x < \infty$. If

$$(5.5) \qquad f(.,\theta) = \int f_1(x) \prod_{i=1}^{n} f_2(x_i,\theta_i)dx, \quad 0 \leq x_i < \infty$$ for every i exists, then it is Schur-concave (convex) function.

In fact, relations (5.4) and (5.5) look different but coincide when $f_2(x,\theta)$ is of the form $\prod_{i=1}^{n} f(x_i-\theta_i)$ otherwise none of them would imply the other.

If we consider the case when the underlying random variables are independent and identically distributed and they have common marginal density function f_1 say. Thus the joint density function of X is $f(x) = \prod_{i=1}^{n} f_1(x_i)$, and in this case f is Schur-concave (convex) if and only if $\log f_1$ is concave (convex). Therefore, for such random variables Schur-concavity and unimodality are equivalent. There are two other related concepts which have been brought up recently, (i) the concept of positive dependence by mixture (PDM), that is, the class of functions which can be represented by a mixture of densities of exchangeable and independent random variables, Shaked (1977), and (ii) the concept of decreasing in transposition (DT) class of functions, that is, those functions which decrease by rearrangement of the components of the random vector, see Hollander et al. (1977). It is expected from the structure of these two classes to be closed under most of the functional operations studied earlier and some other kernel-type transformations.

6. CONCLUDING REMARKS AND OPEN PROBLEMS

The main purpose of statistical methods is to give the users a better inference for the considered problems and this motive lies behind most of the work of statistical and probablistic theories and their applications. It is nice if one is able to have statistics which are sufficient, asymptotically normal, with some invariant property, complete etc., or tests of hypothesis through test functions which are uniformly most powerful unbiased tests etc. Thus, the different schools of thought in statistics such as the classical, Bayesian, Subjectivists, and nonparametricians are mainly different in stressing either robustness or the efficiency of the statistical models.

For example one may try to approximate the actual model by another one which is effectively very close to the actual model, say by using the contiguity approach in order to attain a very high degree of robustness. These approximations are of exponential or infinitely divisible structures. It is shown by the authors (1977a) that sub-classes of infinitely divisible class such as symmetric stable, stable, symmetric L-functions and some non-symmetric L-functions are unimodal. Further, strong unimodality and universality which are essentially similar, are shown to play and important role in plausibility inference, see Barndorff-Nielsen (1976), which is complimentary to the likelihood inference. Also, PT and MLR functions are well behaved functions and very much utilized in decision problems and testing hypotheses, for example see Lehmann (1959 p. 68), where a test function is constructed in terms of some other function g, say, such that the density function is MLR in g. One may see the importance of IFR and other related classes being closed under some functional operation such as convolution and mixing etc. which mostly arise from the practical considerations. The exchangeability structure and other related classes such as Schur-functions, PDM and DT cover, surprisingly, large families of distributions. Such families are shown to arise mainly and in particular in distribution-free tests, since exchangeable hypotheses are all distribution-free. These families may appear in constructing asymptotically optimal distribution-free tests. The role of exchangeability concepts in rank order statistics and other stochastic inference will be investigated by the authors in a forthcoming paper, whereas other related applications can be found in Ahmad and Abouammoh (1977b), Shaked (1977), Proschan and Sethuraman (1977) and Hollander et al. (1977).

Finally, we conclude by the following unsolved problems. Is it true that the convolution of the multivariate strongly unimodal distribution with any multivariate unimodal implies (the same) multivariate unimodality? What are the conditions other than (3.4) and (3.5) under which PT class is closed under convolution? Which are the classes in theorem 4.1, except those already shown, that are closed under the investigated functional operations? Which practicable subclasses of exchangeable family are closed under such functional operations?

REFERENCES

1. A-Hameed, M.S. and Proschan, F. (1973): Nonstationary shock models. Stoch. Proc. Appl., 1, 383-404.

2. Ahmad, R. (1974): On the structure of symmetric sample testing: a distribution-free approach. Ann. Inst. Statist. Math., 26, 233-245.

3. Ahmad, R. (1975): Some characterizations of the exchangeable processes and distribution-free tests in Statistical Distributions in Scientific Work (eds. G.P. Patil et al.), Vol. 3, D. Reidel Pub. Co., Dordrecht, 237-248.

4. Ahmad, R. and Abouammoh, A.M. (1977a): On the structure and applications of infinite divisibility, stability and symmetry in stochastic inference, in Recent Developments in Statistics (eds. J.R. Barra et al.), North-Holland, Amsterdam 1977, 303-317.

5. Ahmad, R. and Abouammoh, A.M. (1977b): On infinite A-divisibility. To appear.

6. Barndorff-Nielsen, O. (1976): Plausibility inference. J. Roy. Statist. Soc. Ser. B, 38, 103-131.

7. Barlow, R.E. and Proschan, F. (1975). Mathematical Theory of Reliability and Life Testing. Holt, Rinehart and Winston.

8. Birnbaum, Z.W., Esary, J.D. and Marshall, A.W. (1966): Stochastic characterization for components and systems. Ann. Math. Statist., 37, 316-325.

9. Block, H.W. and Savits, T.H. (1976): The IFRA closure problem. Ann. prob., 4, 1030-1032.

10. Borell, C. (1975): Convex set in d-space. Period. Math. Hungar., 6, 111-136.

11. Ghurye, S.G. and Wallace, D.L. (1959): A convolution class of Monotone Likelihood ratio families. Ann. Math. Statist., 30, 1158-1164.

12. Hardy, G.H., Littlewood, C.E. and Polya, A. (1952): Inequalities. 2nd ed., Cambridge Univ. press. Cambridge.

13. Hewitt, E. and Savage, L.J. (1955): Symmetric measures on Cartesian products. Trans. Amer. Math. Soc., 80, 470-501.

14. Hollander, M., Proschan, F. and Sethuraman, J. (1977): Functional decreasing in transposition and their applications in ranking problems. Ann. Statist., 5, 722-734.

15. Johnson, N.L. and Kotz, S. (1972): Continuous Multivariate Distributions. -2. Houghton Mifflin Co.

16. Kanter, M. (1977): Unimodality and dominance for symmetric random vectors. Trans. Amer. Math. Soc., 227, 65-85.

17. Karlin, S. (1956): Decision theory for Polya type distributions, case of two actions, I. Proceeding of the third Berkeley symposium on Prob. and Statist., Vol. 1, Univ. of California press, 115-129. (1957): Polyá type distributions, Ann. Math. Statist., 23, 231-308.

18. Lehmann, E.L. (1959): Testing Statistical Hypotheses. Wiley, New York.

19. Lukacs, E. (1970). Characteristic Functions. 2nd ed. Griffin, London.

20. Mudholkar, G.S. (1966): The integral of an invariant unimodal function over an invariant convex set - an inequality and applications. Proc. Amer. Math. Soc., 17, 1327-1333.

21. Marshall, A.W. and Olkin, I. (1974): Majorization in Multivariate distributions. Ann. Statist., 2, 1189-1200.

22. Parthasarathy, K.R. (1967): Probability Measures on Metric Spaces. Academic Press. New York.

23. Prékopa, A. (1973): On logarithmic concave measures and functions. Acta Sci. Math. (Szeged), 34, 335-343.

24. Proschan, F. and Sethuraman, J. (1977): Schur functions in statistics I. The preservation theorem. Ann. Statist., 5, 256-262.

25. Rinott, Y. (1976): On convexity of measures. Ann. Prob., 4, 1020-1026.

26. Schur, I. (1923): Ubereine Klasse non Mittelbildungenmit Anmendungen auf die Determinantentheorie. Sitzber. Berl. Math. Ges, 22, 9-20.

27. Shaked, M. (1977): A concept of positive dependence for exchangeable random variables. Ann. Statist., 5, 505-515.

R. Ahmad and A.M. Abouammoh
University of Strathclyde
Department of Mathematics
26 Richmond Street
Glasgow G1 1XH
Scotland.