

STICHTING  
MATHEMATISCH CENTRUM  
2e BOERHAAVESTRAAT 49  
AMSTERDAM

SP 106

---

(S 387)

W. R. van Zwet

On convexity preserving families of probability distribution

(Statistica Neerlandica, 22 (1968), p. 23-32)



# On convexity preserving families of probability distributions

by W. R. VAN ZWET \*

## 1. Introduction

Let  $F_\theta$  be a family of probability distribution functions on  $R^1$  with parameter  $\theta \in T \subseteq R^1$ , and let  $X$  denote the union of the supports of these distributions. For  $k \geq 0$ , let  $\{g_0, g_1, \dots, g_{k+1}\}$  be a set of real-valued finite functions on  $X$  that are integrable with respect to  $F_\theta$  for all  $\theta \in T$  and define

$$\chi_i(\theta) = \int g_i(x) dF_\theta(x), \quad i = 0, 1, \dots, k+1. \quad (1.1)$$

Following S. KARLIN and W. J. STUDDEN in [3] with a minor modification, we shall say that  $\{g_0, g_1, \dots, g_{k+1}\}$  constitute a *weak complete Tchebycheff system* (WCT-system) if for each  $0 \leq m \leq k+1$  and all  $x_0 < x_1 < \dots < x_m \in X$  the determinant

$$\det [g_i(x_j)]_{i,j=0,\dots,m} \geq 0; \quad (1.2)$$

the system is called a *complete Tchebycheff system* (CT-system) if the inequality is always strict. The difference between this definition of a WCT-system and the one given in [3] is that we retain the case where  $g_0, g_1, \dots, g_m$  are linearly dependent on  $X$  for some  $m \leq k+1$ ; in that case any choice of  $g_{m+1}, \dots, g_{k+1}$  will trivially satisfy definition (1.2). We shall also express inequalities (1.2) by saying that  $g_{k+1}$  is *generalized convex* with respect to the WCT-system  $\{g_0, \dots, g_k\}$ .

Our discussion of WCT-systems will involve the related concept of total positivity (cf. [1]). A function  $f(x, \theta)$  on  $X \times T$  is said to be *totally positive of order  $n$*  ( $TP_n$ ) if for every  $1 \leq m \leq n$ , all  $x_1 < x_2 < \dots < x_m \in X$  and all  $\theta_1 < \theta_2 < \dots < \theta_m \in T$ ,

$$\det [f(x_i, \theta_j)]_{i,j=1,\dots,m} \geq 0. \quad (1.3)$$

The first question that comes to mind in this context is whether one can find conditions on the family  $F_\theta$  that ensure that  $\{\chi_0, \chi_1, \dots, \chi_{k+1}\}$  will be a WCT-system on  $T$  whenever  $\{g_0, g_1, \dots, g_{k+1}\}$  constitutes a WCT-system on  $X$ . If the family  $F_\theta$  possesses densities  $p(x, \theta)$  with respect to a  $\sigma$ -finite measure  $\mu$  with spectrum  $X$  and hence

$$\chi_i(\theta) = \int g_i(x) p(x, \theta) d\mu(x),$$

this question is easily answered. We have for each  $0 \leq m \leq k+1$  (cf. [1])

$$\det [\chi_i(\theta_j)] = \int \dots \int_{x_0 < x_1 < \dots < x_m} \det [g_i(x_j)] \det [p(x_i, \theta_j)] d\mu(x_0) \dots d\mu(x_m)$$

where in each determinant  $i$  and  $j$  run from 0 to  $m$ . It follows that the condition that  $p$  is  $TP_{k+2}$  is certainly sufficient; since we require that  $\{\chi_0, \dots, \chi_{k+1}\}$  will inherit the WCT-property for every WCT-system  $\{g_0, \dots, g_{k+1}\}$ , the condition is essentially also

---

This research was supported in part by the Office of Naval Research Contract Nonr-3656 (18) while the author was visiting at the University of California in Berkeley. Report S 387 (SP 106) of the Statistical Department of the Mathematical Centre at Amsterdam.

\* University of Leiden



necessary (by “essentially” is meant that for any  $\theta_1 < \dots < \theta_m$  the defining inequality (1.3) need not hold on a set of product-measure 0). We note that the fact that  $F_\theta$  are probability distribution functions is not used in establishing the condition.

In view of this general result it is hardly surprising that recent discussions of convexity preserving properties (cf. [1] and [2]) have been confined to families of densities that are totally positive of the appropriate order. However, one usually does not discuss the class of all WCT-systems of a given order but restricts attention to a relatively small subclass (e.g. the case where  $g_i = g_1^i$  for  $i = 0, 1, \dots, k$ ). Also one often imposes additional restrictions on the family  $F_\theta$  to ensure that for those systems  $\{g_0, \dots, g_{k+1}\}$  that are considered,  $\{\chi_0, \dots, \chi_{k+1}\}$  will also belong to some restricted class.

In sections 3 and 4 of this paper we investigate how far the  $TP_{k+2}$  condition for  $p$  can be relaxed in two such restricted cases that seem to be important in practice. Like section 1, the second section is of an expository character.

## 2. Convexity of order $k$

Let  $f$  be a real-valued finite function defined on an arbitrary set  $Y \subseteq R^1$ . For  $k \geq 0$  we shall say that  $f$  is *convex of order  $k$*  ( $C_k$ ) if  $f$  is generalized convex with respect to the CT-system  $\{1, y, y^2, \dots, y^k\}$ , i.e. if for all  $y_1 < y_2 < \dots < y_{k+2}$

$$D_f(y_1, \dots, y_{k+2}) = \begin{vmatrix} 1 & y_1 & y_1^2 & \dots & y_1^k & f(y_1) \\ 1 & y_2 & y_2^2 & \dots & y_2^k & f(y_2) \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & y_{k+2} & y_{k+2}^2 & \dots & y_{k+2}^k & f(y_{k+2}) \end{vmatrix} \geq 0. \quad (2.1)$$

For  $k = 0, 1$ , (2.1) reduces to the ordinary definitions of non-decreasing or (measurable) convex functions. Generally speaking (2.1) is an extension of the concept of non-negative  $(k+1)$ -th derivative.  $C_k$  functions were extensively studied by T. POPOVICIU in [6]. We note that S. KARLIN [1] refers to  $C_k$  functions as convex of order  $(k+1)$ .

If  $P_m$  denotes a polynomial of degree at most  $m$ , then equivalent definitions of the  $C_k$  property are obviously:

- (A) (cf. [1]). For every  $P_k, f - P_k$  changes sign at most  $(k+1)$  times on  $Y$ . If it does have  $(k+1)$  sign-changes, the signs occur in the order  $(-)^{k+1}, (-)^k, \dots, +, -, +$  for increasing values of the argument.
- (B) For every  $y_1 < y_2 < \dots < y_{k+2} \in Y$ , the  $P_{k+1}$  having  $P_{k+1}(y_i) = f(y_i)$ ,  $i = 1, 2, \dots, k+2$ , has non-negative coefficient for its  $(k+1)$ -th degree term.

There is also a close connection with differences. Let

$$\begin{aligned} \Delta_h^1 f(y) &= f(y+h) - f(y) \\ \Delta_h^m f(y) &= \Delta_h^1 \Delta_h^{m-1} f(y) = \sum_{j=0}^m (-1)^{m-j} \binom{m}{j} f(y+jh) \end{aligned} \quad (2.2)$$

and generally

$$\Delta_{h_1, \dots, h_m}^m f(y) = \Delta_{h_m}^1 \Delta_{h_1, \dots, h_{m-1}}^{m-1} f(y) = \sum_{j=0}^m (-1)^{m-j} \sum_{1 \leq i_1 < i_2 < \dots < i_j \leq m} f(y + \sum_{v=1}^j h_{i_v}). \quad (2.3)$$

Furthermore let

$$D_f^*(y_1, \dots, y_{k+2}) = \frac{D_f(y_1, \dots, y_{k+2})}{\prod_{1 \leq i < j \leq k+2} (y_j - y_i)}; \quad (2.4)$$

since the denominator is positive for  $y_1 < y_2 < \dots < y_{k+2}$ ,  $D_f$  may be replaced by  $D_f^*$  in definition (1.1). The following relation between  $D_f^*$  and differences may be proved by induction on  $k$ .

### Lemma 2.1

If  $\Pi$  denotes the set of permutations  $\pi = [\pi(1), \pi(2), \dots, \pi(k+1)]$  of the numbers  $1, 2, \dots, k+1$ , then

$$\Delta_{h_1, \dots, h_{k+1}}^{k+1} f(y) = \prod_{i=1}^{k+1} h_i \sum_{\pi \in \Pi} D_f^* \left( y, y + h_{\pi(1)}, \dots, y + \sum_{v=1}^{k+1} h_{\pi(v)} \right). \quad (2.5)$$

We note that for  $h_1 = h_2 = \dots = h_{k+1} = h$ , (2.5) reduces to

$$\Delta_h^{k+1} f(y) = (k+1)! h^{k+1} D_f^*[y, y+h, \dots, y+(k+1)h]. \quad (2.6)$$

It follows from lemma 2.1 that if  $f$  is  $C_k$  on  $Y$ , then for all  $h_1, h_2, \dots, h_{k+1} > 0$ ,

$$\Delta_{h_1, \dots, h_{k+1}}^{k+1} f(y) \geq 0 \quad (2.7)$$

whenever defined, i.e. whenever all  $y + \sum_{v=1}^j h_{i_v} \in Y$ .

In the special case that  $Y$  is an interval there is also a converse result and the following definition of the  $C_k$  property is equivalent to (2.1) in this case:

(C)  $f$  is (Lebesgue)-measurable and for  $h > 0$ ,  $y \in Y$ ,  $y + (k+1)h \in Y$ ,

$$\Delta_h^{k+1} f(y) \geq 0. \quad (2.8)$$

In this case, however, the  $C_k$  property is hardly a generalization of non-negative  $(k+1)$ -th derivative at all. In fact, if  $Y$  is an open interval and  $k \geq 1$ , definition (2.1) ensures continuity of  $f$  on  $Y$  and is equivalent to

(D)  $f$  is  $(k-1)$  times continuously differentiable and  $f^{(k-1)}$  is convex on  $Y$ .

Finally we consider the special case where  $Y$  is a set of consecutive integers. For integer  $h > 0$

$$\Delta_h^{k+1} f(y) = \sum_{h_1=0}^{h-1} \dots \sum_{h_{k+1}=0}^{h-1} \Delta_1^{k+1} f \left( y + \sum_{j=1}^{k+1} h_j \right). \quad (2.9)$$

Combining (2.6) and (2.9) we find that the  $C_k$  property may be defined in this case by

(E) For all  $y$ ,  $y+k+1 \in Y$

$$\Delta_1^{k+1} f(y) \geq 0. \quad (2.10)$$



For further details concerning the definitions given above the reader is referred to [6].

Let  $f_1$  and  $f_2$  be real-valued finite functions on  $Y$ . We shall say that  $f_2$  is  $C_k$  with respect to  $f_1$  on  $Y$  if there exists a  $C_k$  function  $f$  on  $f_1(Y)$  such that  $f_2 = f(f_1)$  on  $Y$ . If  $f_1$  is non-decreasing on  $Y$  and  $f_2$  is constant on any set where  $f_1$  is constant, this reduces to

$$\begin{vmatrix} 1 & f_1(y_1) & f_1^2(y_1) & \dots & f_1^k(y_1) & f_2(y_1) \\ 1 & f_1(y_2) & f_1^2(y_2) & \dots & f_1^k(y_2) & f_2(y_2) \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & f_1(y_{k+2}) & f_1^2(y_{k+2}) & \dots & f_1^k(y_{k+2}) & f_2(y_{k+2}) \end{vmatrix} \geq 0 \quad (2.11)$$

for all  $y_1 < y_2 < \dots < y_{k+2} \in Y$ .

### 3. Preserving convexity of order $k$

Returning to the setup of section 1, we let  $g$  be a real-valued finite function on  $X$  that is integrable with respect to  $F_\theta$  for all  $\theta \in T$  and define

$$\chi(\theta) = \int g(x) dF_\theta(x).$$

We shall say that the family  $F_\theta$  preserves convexity of order  $k$  if  $\chi$  is  $C_k$  on  $T$  whenever  $g$  is  $C_k$  on  $X$ , i.e. whenever  $g$  is generalized convex with respect to  $\{1, x, \dots, x^k\}$  then  $\chi$  is generalized convex with respect to  $\{1, \theta, \dots, \theta^k\}$ . In [1] S. KARLIN has shown that if densities  $p(x, \theta)$  with respect to  $\mu$  exist, then a sufficient condition for  $F_\theta$  to preserve convexity of order  $k$  is that  $p$  is  $TP_{k+2}$  and that whenever  $g$  is a polynomial of exact degree  $m \leq k$ , then  $\chi$  is also a polynomial of exact degree  $m$ . According to the result of section 1 the first part of this condition ensures that  $\chi$  is generalized convex with respect to the WCT-system

$$\int x^i dF_\theta(x), \quad i = 0, 1, \dots, k,$$

whereas the second part ensures that this is equivalent to generalized convexity with respect to  $\{1, \theta, \dots, \theta^k\}$ .

However, this condition is not necessary. For  $k = 0$  a condition that is necessary as well as sufficient was given by J. KRZYZ in [4]:

#### Lemma 3.1

$\chi$  is non-decreasing on  $T$  whenever  $g$  is non-decreasing on  $X$  if and only if the family  $F_\theta$  is stochastically increasing (i.e.  $F_\theta(x)$  is non-increasing in  $\theta$  for every fixed  $x$ ).

Since the  $TP_2$  property of  $p$  is equivalent to monotone likelihood ratio, KRZYZ's condition is weaker than KARLIN's for  $k = 0$  (cf. [5]).

For general  $k$  it is also easy to find a necessary and sufficient condition, provided that we restrict attention to those  $C_k$  functions  $g$  that can be extended to a  $C_k$  function on an open interval containing  $X$ . Since the convex functions constitute a convex cone spanned by the linear functions and functions of the form



$$h(x) = \begin{cases} 0 & \text{for } x \leq x_0 \\ x - x_0 & \text{for } x > x_0, \end{cases}$$

we find from definition  $D$  of section 2 that the convex cone of  $C_k$  functions is spanned by the polynomials  $P_k$  of degree at most  $k$  and functions of the form

$$h_k(x) = \begin{cases} 0 & \text{for } x \leq x_0 \\ (x - x_0)^k & \text{for } x > x_0. \end{cases}$$

For  $k = 0$  this is obviously also true. It follows that it is sufficient as well as necessary to require that  $\chi$  be  $C_k$  whenever  $g$  is of one of the forms mentioned above. However, if  $g$  is a polynomial of degree at most  $k$ , then so is  $-g$  and as a result both  $\chi$  and  $-\chi$  are required to be  $C_k$ , which implies that  $\chi$  is also a polynomial of degree at most  $k$ . Hence we have proved

### Lemma 3.2

$\chi$  is  $C_k$  on  $T$  whenever  $g$  is  $C_k$  on an open interval containing  $X$ , if and only if for every  $x_0$

$$\int (x - x_0)^k dF_\theta(x)$$

is  $C_k$  on  $T$  and whenever  $g$  is a polynomial of degree at most  $k$  the same holds for  $\chi$ .

We note that for  $k \leq 1$  the condition that the  $C_k$  function  $g$  can be extended to a  $C_k$  function on an open interval containing  $X$  is always satisfied. For  $k = 0$  the lemma reduces to lemma 3.1.

Although for  $k \geq 1$  lemma 3.2 seems to be fairly useless for practical purposes, the results obtained so far do seem to indicate that there exists a large class of  $C_k$  preserving families that do not possess any total positivity properties. The results in the remainder of this section exhibit a number of these families.

### Theorem 3.1

Let  $F_0$  and  $F$  be distribution functions with characteristic functions  $\varphi_0$  and  $\varphi$  respectively, and suppose that  $F$  is infinitely divisible and has  $F(-0) = 0$ . If for  $t \geq 0$ ,  $F_t$  denotes the distribution function corresponding to  $\varphi_0 \cdot \varphi^t$ , then the family  $F_t$ ,  $0 \leq t < \infty$ , preserves convexity of all orders.

### Proof

Let  $G_t$  denote the distribution function corresponding to  $\varphi^t$  and let  $X_t$  ( $t \geq 0$ ) be a stochastic process with non-negative stationary independent increments for which  $X_0$ ,  $X_{s+t} - X_s$  and  $X_t$  ( $s, t \geq 0$ ) have distribution functions  $F_0$ ,  $G_t$  and  $F_t$  respectively.

For fixed  $t \geq 0$  and  $h > 0$  define

$$Z_i = X_{t+ih} - X_{t+(i-1)h}, \quad i = 1, 2, \dots, k+1.$$

$Z_1, Z_2, \dots, Z_{k+1}$  are independent and identically distributed random variables that are also independent of  $X_t$ . Hence, because of the exchangeability of  $Z_1, \dots, Z_{k+1}$ ,

$$\begin{aligned}
& E \left[ \sum_{j=0}^{k+1} (-1)^{k+1-j} \binom{k+1}{j} g(X_{t+jh}) \mid X_t = x \right] = \\
& = E \left[ \sum_{j=0}^{k+1} (-1)^{k+1-j} \binom{k+1}{j} g(x + Z_1 + \dots + Z_j) \right] = \\
& = E \left[ \sum_{j=0}^{k+1} (-1)^{k+1-j} \sum_{1 \leq i_1 < \dots < i_j \leq k+1} g(x + Z_{i_1} + \dots + Z_{i_j}) \right] = \\
& = E [\Delta_{Z_1, \dots, Z_{k+1}}^{k+1} g(x)].
\end{aligned}$$

Since  $Z_1, \dots, Z_{k+1} \geq 0$  with probability 1, the last expression is non-negative for every  $C_k$  function  $g$  and all  $x$  by (2.7). As a result

$$\Delta_h^{k+1} \chi(t) = E \left[ \sum_{j=0}^{k+1} (-1)^{k+1-j} \binom{k+1}{j} g(X_{t+jh}) \right] \geq 0$$

for all  $t \geq 0$  and  $h > 0$ . As  $\chi$  is a measurable function defined on the interval  $[0, \infty)$ , it is  $C_k$  by definition C of section 2.

If we consider only integer values of  $t$  in theorem 3.1, we may drop the assumption that  $F$  is infinitely divisible without affecting the proof. The  $C_k$  character of  $\chi$  on the integers now follows from  $\Delta_1^{k+1} \chi \geq 0$  by definition E of section 2. Specializing to the case where  $F_0$  is degenerate at 0 we obtain

### Corollary 3.1

Every family  $F_n$ ,  $n = 1, 2, \dots$ , of  $n$ -fold convolutions of a distribution function  $F_1$  having  $F_1(-0) = 0$  preserves convexity of every order.

We note that the fact that  $F_n$  preserves convexity of order  $k$  was proved by S. KARLIN and F. PROSCHAN in [2] under the additional assumption that  $F_1$  possesses a density  $p$  that is a Pólya frequency density of order  $k+2$  (i.e.  $p(x-y)$  is  $TP_{k+2}$  in  $x$  and  $y$ ).

Another special case of theorem 3.1 is obtained by assuming  $F$  to be degenerate at 1, in which case the theorem reduces to:

Every location parameter family  $F_\theta(x) = G(x-\theta)$ ,  $-\infty < \theta < \infty$ , preserves convexity of every order.

This result is of course rather trivial. Without invoking theorem 3.1, it follows at once from

$$\Delta_h^{k+1} \chi(\theta) = \Delta_h^{k+1} \int g(x+\theta) dG(x) = \int \Delta_h^{k+1} g(x+\theta) dG(x).$$

In the same manner one easily verifies:

Every scale parameter family  $F_\theta(x) = G(x/\theta)$ ,  $0 < \theta < \infty$ , preserves convexity of every odd order. If moreover  $G(-0) = 0$ , then the family preserves convexity of all orders.

## 4. Invariant convexity preserving families

Let  $g_1, g_2, \chi_1$  and  $\chi_2$  be defined as in section 1. We shall say that the family  $F_\theta$  is invariant convexity preserving if, whenever  $g_1$  is non-decreasing and  $g_2$  is convex with



respect to  $g_1$  on  $X$ , then  $\chi_1$  is non-decreasing and  $\chi_2$  is convex with respect to  $\chi_1$  on  $T$ . In terms of WCT-systems we may express this property by requiring that for every WCT-system of the form  $\{1, g_1, g_2\}$  the corresponding system  $\{1, \chi_1, \chi_2\}$  is also a WCT-system.

In the first place this definition asserts that the family  $F_\theta$  preserves the monotonicity of  $g_1$  and hence by lemma 1 the family is stochastically increasing;  $F_\theta$  also preserves convexity (of order 1) provided that the parameter is subjected to a suitable non-decreasing transformation

$$\eta = \eta(\theta) = \int x dF_\theta(x).$$

Moreover, this convexity preserving property is invariant under non-decreasing transformations  $g_1$  of the random variable, the appropriate monotone transformation of  $\theta$  then becoming  $\chi_1$ . It is precisely because of this invariance that we do not require that  $F_\theta$  be convexity preserving with respect to  $\theta$  itself, i.e. that  $\eta$  be linear in  $\theta$ . This property would be destroyed by non-linear transformations  $g_1$  anyway and would only result in fixing a possibly awkward parametrization.

From the general result of section 1 it follows that  $F_\theta$  is invariant convexity preserving if the density  $p$  is  $TP_3$ . The following theorem provides a necessary and sufficient condition.

#### Theorem 4.1

Define  $\bar{F}_\theta(x) = 1 - F_\theta(x)$ . The family  $F_\theta$  is invariant convexity preserving if and only if  $\{1, \bar{F}_\theta(x_1), \bar{F}_\theta(x_2)\}$  is a WCT-system on  $T$  for every fixed pair  $x_1 < x_2$ .

#### Proof

The condition asserts that for  $x_1 < x_2$  and  $\theta_0 < \theta_1 < \theta_2$ ,

$$\begin{vmatrix} 1 & \bar{F}_{\theta_0}(x_1) \\ 1 & \bar{F}_{\theta_1}(x_1) \end{vmatrix} \geq 0, \quad \begin{vmatrix} 1 & \bar{F}_{\theta_0}(x_1) & \bar{F}_{\theta_0}(x_2) \\ 1 & \bar{F}_{\theta_1}(x_1) & \bar{F}_{\theta_1}(x_2) \\ 1 & \bar{F}_{\theta_2}(x_1) & \bar{F}_{\theta_2}(x_2) \end{vmatrix} \geq 0. \quad (4.1)$$

The first inequality means that  $F_\theta$  is stochastically increasing and we have already remarked that this is necessary and sufficient for  $\chi_1$  to be non-decreasing whenever  $g_1$  is. We may therefore assume that  $\bar{F}_\theta(x)$  is non-decreasing in  $\theta$  for every fixed  $x$  and restrict attention to the second inequality.

Let  $g_1$  be non-decreasing and let  $g_2 = f(g_1)$  where  $f$  is convex on  $g_1(X)$ . Since a convex function can be extended to a convex function on an interval, the same reasoning that we used in the proof of lemma 3.2 shows that we need only be concerned with functions  $f$  that are linear and functions  $f$  of the form

$$f(y) = \begin{cases} 0 & \text{for } y \leq y_0 \\ y - y_0 & \text{for } y > y_0. \end{cases} \quad (4.2)$$

Without loss of generality we may assume that  $y_0 = g_1(x_0) \in g_1(X)$ . For linear  $f$ ,  $\chi_2$  is linear and hence convex with respect to  $\chi_1$ . Only functions  $f$  of the form (4.2)



remain to be considered and as a result we have the following necessary and sufficient condition for a stochastically increasing family  $F_\theta$  to be invariant convexity preserving:

For every non-decreasing  $g_1$  and every  $x_0 \in X$ ,

$$\chi_2(\theta) = \int_{x_0}^{\infty} [g_1(x) - g_1(x_0)] dF_\theta(x)$$

is convex with respect to  $\chi_1(\theta)$ .

By an approximation argument one shows that it is sufficient to consider only those functions  $g_1$  that are left-continuous, non-decreasing step-functions assuming finitely many values. But then the above condition becomes:

For all  $m = 1, 2, \dots$ , all  $x_1 < x_2 < \dots < x_m$ , all  $\alpha_i > 0$ ,  $i = 1, 2, \dots, m$ , all  $1 \leq i_0 \leq m$  and all  $c$ ,

$$\sum_{i=i_0}^m \alpha_i \bar{F}_\theta(x_i) \quad (4.3)$$

is convex with respect to

$$\sum_{i=1}^m \alpha_i \bar{F}_\theta(x_i) + c. \quad (4.4)$$

Since (4.4) is non-decreasing in  $\theta$  and (4.3) is constant on any set where (4.4) is constant, the determinantal convexity definition (2.11) for  $k = 1$  applies. By subtracting from the second column in this determinant we find that convexity of (4.3) with respect to (4.4) is equivalent to

$$\begin{vmatrix} 1 & \sum_{i=1}^{i_0-1} \alpha_i \bar{F}_{\theta_0}(x_i) & \sum_{i=i_0}^m \alpha_i \bar{F}_{\theta_0}(x_i) \\ \vdots & \vdots & \vdots \\ 1 & \sum_{i=1}^{i_0-1} \alpha_i \bar{F}_{\theta_2}(x_i) & \sum_{i=i_0}^m \alpha_i \bar{F}_{\theta_2}(x_i) \end{vmatrix} = \quad (4.5)$$

$$= \sum_{i=1}^{i_0-1} \sum_{j=i_0}^m \alpha_i \alpha_j \begin{vmatrix} 1 & \bar{F}_{\theta_0}(x_i) & \bar{F}_{\theta_0}(x_j) \\ \vdots & \vdots & \vdots \\ 1 & \bar{F}_{\theta_2}(x_i) & \bar{F}_{\theta_2}(x_j) \end{vmatrix} \geq 0.$$

By choosing  $i_0 = m = 2$  we find that condition (4.1) is necessary; since every term in (4.5) has  $x_i < x_j$  it is also sufficient. This completes the proof of the theorem.

It may be of interest to compare the sufficient condition that  $F_\theta$  possesses a  $TP_3$  density  $p(x, \theta)$  with the necessary and sufficient condition of the theorem. One easily shows that the  $TP_3$  assumption for  $p$  implies that  $\bar{F}_\theta(x)$  is  $TP_3$ , or



$$\begin{vmatrix} \bar{F}_{\theta_0}(x_0) & \bar{F}_{\theta_0}(x_1) \\ \bar{F}_{\theta_1}(x_0) & \bar{F}_{\theta_1}(x_1) \end{vmatrix} \geq 0, \quad \begin{vmatrix} \bar{F}_{\theta_0}(x_0) & \bar{F}_{\theta_0}(x_1) & \bar{F}_{\theta_0}(x_2) \\ \bar{F}_{\theta_1}(x_0) & \bar{F}_{\theta_1}(x_1) & \bar{F}_{\theta_1}(x_2) \\ \bar{F}_{\theta_2}(x_0) & \bar{F}_{\theta_2}(x_1) & \bar{F}_{\theta_2}(x_2) \end{vmatrix} \geq 0, \quad (4.6)$$

for  $x_0 < x_1 < x_2$  and  $\theta_0 < \theta_1 < \theta_2$ . By letting  $x_0$  tend to  $-\infty$  we see that (4.6) implies (4.1). Hence the condition that  $\bar{F}_\theta(x)$  be  $TP_3$  is also sufficient for  $F_\theta$  to be invariant convexity preserving.

If we restrict ourselves to the special case where the parameter set  $T$  is an interval and  $\bar{F}_\theta(x)$  is differentiable with respect to  $\theta$ , it turns out that theorem 4.1 involves a  $TP_2$  instead of a  $TP_3$  condition.

#### Theorem 4.2

Let  $T$  be an interval and let  $q(x, \theta) = (\partial/\partial\theta)\bar{F}_\theta(x)$  be defined on  $T$  for all  $x$ . Then the family  $F_\theta$  is invariant convexity preserving if and only if  $q$  is  $TP_2$ .

#### Proof

The first inequality in (4.1) is equivalent to  $q \geq 0$ . Since  $\bar{F}_\theta(x_2)$  is constant on any set where  $\bar{F}_\theta(x_1) + \bar{F}_\theta(x_2)$  is constant and the latter is non-decreasing in  $\theta$ , the second inequality of (4.1) asserts that  $\bar{F}_\theta(x_2)$  is convex with respect to  $\bar{F}_\theta(x_1) + \bar{F}_\theta(x_2)$ . This in turn is equivalent to  $q(x_1, \theta_1)q(x_2, \theta_2) - q(x_1, \theta_2)q(x_2, \theta_1) \geq 0$  for  $x_1 < x_2$  and  $\theta_1 < \theta_2$ .

It is tempting to ask whether theorem 4.2 can be generalized. One conceivable generalization would deal with invariant  $C_k$  preserving families  $F_\theta$ , i.e. families for which  $\chi_1$  is non-decreasing and  $\chi_2$  is  $C_k$  with respect to  $\chi_1$  whenever  $g_1$  is non-decreasing and  $g_2$  is  $C_k$  with respect to  $g_1$ . However, even a cursory inspection shows that only trivial examples of such families exist. The necessary requirement that  $\chi_2$  be a polynomial in  $\chi_1$  of degree at most  $k$  whenever  $g_2$  is a polynomial in  $g_1$  of degree at most  $k$ , can not be satisfied for every non-decreasing  $g_1$  except in a trivial manner.

A more promising generalization is to consider families  $F_\theta$  that transform WCT-systems  $\{1, g_1, \dots, g_{k+1}\}$  into WCT-systems  $\{1, \chi_1, \dots, \chi_{k+1}\}$ . If one restricts attention to the case where  $X$  and  $T$  are intervals and  $g_i$  and  $F_\theta$  satisfy certain regularity conditions, one shows in a fairly straightforward manner that a necessary and sufficient condition on  $F_\theta$  is that  $q$  be  $TP_{k+1}$ , thus generalizing lemma 3.1 and theorem 4.2 to the case where  $k \geq 2$ . We may conclude that although something may be lost for  $k \geq 2$ , the basic reason that theorems 4.1 and 4.2 work is not the fact that  $k = 1$  in that case, but that  $g_0 \equiv 1$  and that  $F_\theta$  are probability distribution functions.



## References

- [1] KARLIN, S. (1963), Total positivity and convexity preserving transformations. Convexity: Proc. Symp. Pure Math. 7, 329–347, Amer. Math. Soc., Providence.
- [2] KARLIN, S. and F. PROSCHAN (1960), Pólya type distributions of convolutions. Ann. Math. Statist. **31**, 721–736.
- [3] KARLIN, S. and W. J. STUDDEN (1966), Tchebycheff systems: with applications in analysis and statistics. John Wiley & Sons, New York, London, Sydney.
- [4] KRZYŻ, J. (1952), On monotonicity-preserving transformations. Ann. Univ. Mariae Curie-Skłodowska Sect. A **6**, 91–111.
- [5] LEHMANN, E. L. (1959), Testing statistical hypotheses. John Wiley & Sons, New York.
- [6] POPOVICIU, T. (1945), Les fonctions convexes. Herman et Cie, Paris.