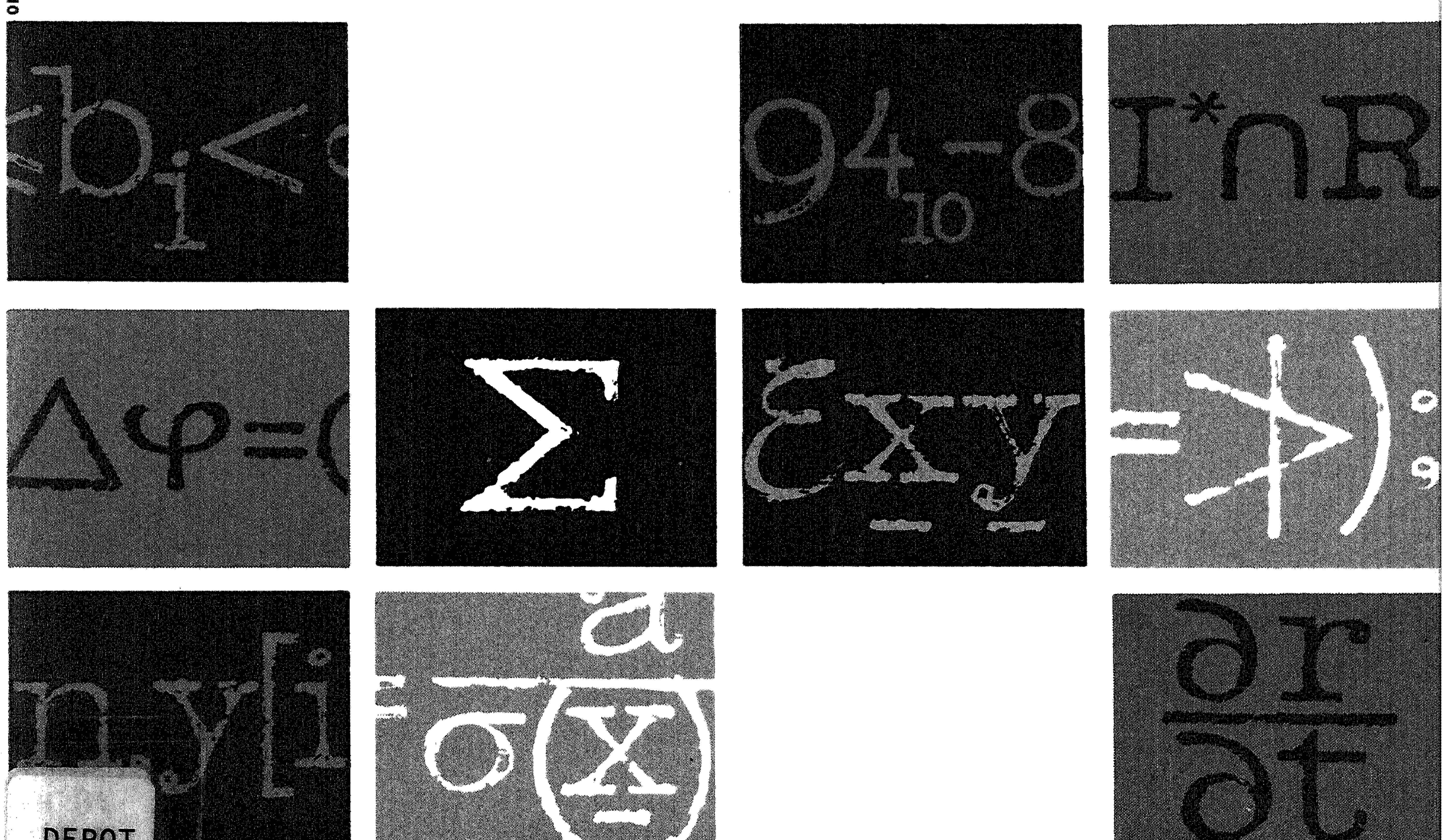
CONVEX TRANSFORMATIONS OF RANDOM VARIABLES

W. R. VAN ZWET

ontwerp jan brein



MATHEMATICAL CENTRE TRACTS 7

MATHEMATICAL CENTRE TRACTS

7

CONVEX TRANSFORMATIONS OF RANDOM VARIABLES

BY

W.R. VAN ZWET

MATHEMATISCH CENTRUM AMSTERDAM

1964

CONTENTS

Chapter 1

INTRODUCTION

There are several reasons to study the behaviour of non-decreasing, convex functions of random variables. Perhaps the most technical and least transparent one was that which originally aroused the author's interest. Suppose that $\underline{x}$ is a real-valued random variable having a continuous distribution function F, and let $\underline{x}_{1:n} < \underline{x}_{2:n} < \ldots < \underline{x}_{n:n}$ be an ordered sample of size n from this distribution. We ask for the expected value $E\,\underline{x}_{i:n}$ of the i-th order statistic and find

$$E\,\underline{x}_{i:n} = \frac{n!}{(i-1)!(n-i)!} \int_0^1 G(y)\, y^{i-1}\, (1-y)^{n-i}\, dy \quad ,$$

where G denotes the inverse of F. Although these quantities are of great interest in the theory of linear estimation and non-parametric statistics, surprisingly little is known about them for small samples. For some well known distributions tables of $E\,\underline{x}_{i:n}$ are available, but analytically only asymptotic results have been obtained. It is well known that if i and n tend to infinity in such a way that $\lim \frac{i}{n} = r$ then generally $\lim F(E\,\underline{x}_{i:n}) = r$ (cf. W. HOEFFDING [16]). Also G. BLOM [4] has refined this result by investigating a second order term from which large sample inequalities may be derived.

Suppose, however, that we have some information - numerical or otherwise - about $F(E\,\underline{x}_{i:n})$ because of e.g. the special character of F, but that we wish to obtain information about $F^{*}(E\,\underline{x}^{*}_{i:n})$ where $\underline{x}^{*}$ is another random variable with distribution function F^{*} with inverse

G^*. Then, if the function G^*F is convex on the smallest interval I with $P(\underline{x} \in I) = 1$, we have by JENSEN's inequality [19]

$$G^*F(E\,\underline{x}_{i:n}) \leqq E\,G^*F(\underline{x}_{i:n}) =$$

$$= \frac{n!}{(i-1)!(n-i)!} \int_0^1 G^*FG(y)\, y^{i-1} (1-y)^{n-i}\, dy =$$

$$= \frac{n!}{(i-1)!(n-i)!} \int_0^1 G^*(y)\, y^{i-1} (1-y)^{n-i}\, dy = E\,\underline{x}^*_{i:n}\ , \text{ or}$$

(1.1) $$F(E\,\underline{x}_{i:n}) \leqq F^*(E\,\underline{x}^*_{i:n}) \quad .$$

By a slight modification of JENSEN's inequality the same result may be proved for $i \geqq \frac{n+1}{2}$ if F and F^* are both symmetric distributions and G^*F is concave-convex on I. It is easily seen that both convexity conditions define weak-order relations for distribution functions; we write $F \underset{c}{<} F^*$ if G^*F is convex on I, and $F \underset{s}{<} F^*$ if both F and F^* are symmetrical and G^*F is concave-convex on I.

One may ask conversely whether the inequalities (1.1) characterize the order relations. It turns out that under quite general conditions the answer is in the affirmative. If i and n tend to infinity and $\lim \frac{i}{n} = r$, $0 < r < 1$, then $F \underset{c}{<} F^*$ if (1.1) holds asymptotically for all $0 < r < 1$, and $F \underset{s}{<} F^*$ if F and F^* are symmetrical and (1.1) holds asymptotically for all $\frac{1}{2} < r < 1$. It is therefore possible to generate small sample inequalities like (1.1) from their large sample counterparts. For symmetric distributions one may obtain explicit bounds for $F(E\,\underline{x}_{i:n})$ of the type

$$F(E\,\underline{x}_{i:n}) \overset{\geq}{\underset{<}{=}} \frac{i-\alpha}{n+1-2\alpha}$$

by s-comparison with the class of so-called symmetric inverse beta distributions. This class can be shown to possess some very special properties that make it particularly well suited as standards for s-comparison.

All this can be done in a purely formal way without reference to the statistical meaning of the function $\phi(x) = G^{*}F(x)$. From $F^{*}\phi(x) = F(x)$, or $P(\underline{x}^{*} \leq \phi(x)) = P(\underline{x} \leq x)$, it is clear, however, that F^{*} is the distribution of $\phi(\underline{x})$ or, put differently, that the random variable $\phi(\underline{x})$ has the same distribution as the random variable $\underline{x}^{*}$. Hence $F \underset{c}{<} F^{*}$ if and only if F^{*} is the distribution of a non-decreasing, convex transform $\phi(\underline{x})$ of $\underline{x}$, whereas $F \underset{s}{<} F^{*}$ means that both distributions are symmetrical and F^{*} is the distribution of a non-decreasing, concave-convex transform $\phi(\underline{x})$ of $\underline{x}$. It appears therefore that we are simply studying properties of non-decreasing, convex and concave-convex functions of random variables.

One intuitively feels that a non-decreasing, convex transform of a random variable increases the skewness to the right and it is interesting to find that the best known measures of skewness, being the standardized odd central moments, react as expected:

$$(1.2) \qquad \frac{\mu_{2k+1}(\underline{x})}{\sigma^{2k+1}(\underline{x})} \leq \frac{\mu_{2k+1}(\phi(\underline{x}))}{\sigma^{2k+1}(\phi(\underline{x}))} = \frac{\mu_{2k+1}(\underline{x}^{*})}{\sigma^{2k+1}(\underline{x}^{*})}, \quad k=1,2,\ldots,$$

if ϕ is non-decreasing, convex. If F and F^{*} are symmetrical and ϕ is non-decreasing, concave-convex, one expects the distribution F^{*} to show heavier tails, and indeed one finds

$$(1.3) \qquad \frac{\mu_{2k}(\underline{x})}{\sigma^{2k}(\underline{x})} \leq \frac{\mu_{2k}(\phi(\underline{x}))}{\sigma^{2k}(\phi(\underline{x}))} = \frac{\mu_{2k}(\underline{x}^{*})}{\sigma^{2k}(\underline{x}^{*})}, \quad k=2,3,\ldots \, .$$

Also it turns out that under quite general conditions one may again characterize both weak-order relations for continuous distributions by requiring (1.2) to hold for large sample order statistics. If i and n tend to infinity and $\lim \frac{i}{n} = r$, $0 < r < 1$, then $F \underset{c}{<} F^{*}$ if (1.2) holds asymptotically for some fixed k, all $0 < r < 1$, and order statistics $\underline{x}_{i:n}$ and $\underline{x}^{*}_{i:n}$; if the same is true for all $\frac{1}{2} < r < 1$ and F and F^{*} are symmetrical, then $F \underset{s}{<} F^{*}$.

The disadvantage of measuring skewness or kurtosis by a single number like third or fourth standardized central moment is obvious. The penalty for imposing these simple orderings on distribution functions is quite naturally that the order relation is often not sufficiently meaningful to be of any use at all. The fact that the debate, about what exactly one does measure by the standardized fourth central moment, has been going on in scientific journals until quite recently, should serve as a warning.

Still, when one compares the statistical properties of two distributions, greater skewness or heavier tails in one distribution do play an important role. The author, however, is inclined to think in terms of the weak-order relations $\underset{c}{<}$ and $\underset{s}{<}$ to indicate this, rather than in terms of moments. The loss of applicability, due to the fact that not every pair of distributions is comparable according to $\underset{c}{<}$ or $\underset{s}{<}$, should be amply compensated in many instances by much more significant results in case the distributions are comparable.

To sustain this optimistic view three problems of comparison of distributions are tackled where skewness and "kurtosis" obviously play an important role. Two of these are concerned with the performance of statistical tests under non-standard conditions, whereas the third deals with a problem in estimation theory. In all three cases the desired result is easily obtained by making use of properties of the relations $\underset{c}{<}$ and $\underset{s}{<}$. Among these results we may mention a theorem on the relative asymptotic efficiency $e_{W,N}(F)$ of WILCOXON's test to the normal scores test when the underlying distribution is given by $F \varepsilon S$. The theorem states that under rather general conditions $F \underset{s}{<} F^{*}$ implies $e_{W,N}(F) \leqq e_{W,N}(F^{*})$.

For various reasons the results given above are discussed in a slightly different order in the following chapters. We start by giving JENSEN's inequality and its modification, as well as the moment inequalities (1.2) and (1.3), in chapter 2. In chapter 3 we introduce order statistics and derive the asymptotic expressions for their expectations and central moments that are needed to establish the characterization theorems of chapter 4. In chapter 4 we consider the relations $\underset{c}{<}$ and $\underset{s}{<}$, restate their properties and prove the converse theorems that characterize the weak orderings in terms of large sample inequalities. Also in chapters 3 and 4 we develop an asymptotic expression for the median of $\underline{x}_{i:n}$, discuss its connection with the problem of plotting on probability papers, and prove another characterization of $\underset{c}{<}$ and $\underset{s}{<}$ in terms of a measure of skewness based on the median. Symmetric inverse beta distributions and the small sample inequalities to be derived from s-comparison with this class are discussed in chapter 5. Applications to hypothesis testing and estimation are treated in chapter 6.

In chapters 4 and 5 a number of examples of the relations $\underset{c}{<}$ and $\underset{s}{<}$ and the resulting inequalities are given. We only mention here that the gamma distributions and the symmetric beta distributions are found to be ordered according to $\underset{c}{<}$ and $\underset{s}{<}$ respectively.

ACKNOWLEDGMENT

The author is endebted to Professors dr. J. HEMELRIJK and dr. J. TH. RUNNENBURG for helpful and challenging discussions.

Chapter 2

MOMENT INEQUALITIES

2.1. PRELIMINARIES

Throughout this study we shall strictly adhere to a number of notational conventions in order to avoid unnecessary repetition in the formulation of theorems and other statements. Most of these conventions, together with some preliminary material, will be introduced in the first section of the chapter where they are first needed. For the present chapter this amounts to the following.

All random variables discussed will be real-valued. We shall distinguish them from numbers and algebraic variables by underlining their symbols, thus e.g. $P(\underline{x} < x)$ will be the probability that the random variable $\underline{x}$ assumes a value smaller than the number x. For the random variable $\underline{x}$, which we shall always suppose to be non-degenerate, we shall denote by I the smallest interval (open, half-open or closed, finite or infinite) for which $P(\underline{x} \in I) = 1$. We define the distribution function F[†)] of $\underline{x}$ by

$$(2.1.1) \qquad F(x) = \tfrac{1}{2}P(\underline{x} < x) + \tfrac{1}{2}P(\underline{x} \leq x) \quad .$$

Hence the interval I is the largest interval for which $0 < F(x) < 1$; it will be called the support of F.

---

†) Whenever this is possible without risking confusion we shall suppress the argument when denoting a function. Thus we write F rather than F(x) to denote the function as opposed to the value it assumes at x.

We denote the expected value of a random variable by the symbol E, thus e.g.

$$E\,\underline{x} = \int_I x\, dF(x) \quad ,$$

where the right-hand side denotes a STIELTJES integral. We shall say that this expectation exists if x is summable with respect to F, thus requiring expectations to be finite. The central and absolute central moments of $\underline{x}$ will be denoted by $\mu_k(\underline{x})$, and $\nu_k(\underline{x})$ respectively, thus

$$\mu_k(\underline{x}) = \int_I (x - E\,\underline{x})^k\, dF(x) \qquad \text{and}$$

$$\nu_k(\underline{x}) = \int_I |x - E\,\underline{x}|^k\, dF(x) \quad ;$$

the variance $\mu_2(\underline{x})$ of $\underline{x}$ will also be written $\sigma^2(\underline{x})$.

A median $m(\underline{x})$ of $\underline{x}$ will be defined by

$$P(\underline{x} \leq m(\underline{x})) \geq \tfrac{1}{2} \quad , \quad P(\underline{x} \geq m(\underline{x})) \geq \tfrac{1}{2} \quad ;$$

we note that a median need not be unique: if $F(x) = \frac{1}{2}$ on some sub-interval of I, then every point of this sub-interval and its end-points are medians of $\underline{x}$.

We shall say that the real-valued function ϕ defined on I is convex on I if for all $x_1, x_2 \in I$ and $0 \leq \lambda \leq 1$

$$\phi(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda\phi(x_1) + (1 - \lambda)\phi(x_2) \quad ,$$

i.e. the graph of ϕ lies on or below any chord; or equivalently, if for all $x_1 < x_0 < x_2 \in I$

$$\frac{\phi(x_2) - \phi(x_0)}{x_2 - x_0} \geq \frac{\phi(x_0) - \phi(x_1)}{x_0 - x_1} \quad ,$$

i.e. for every interior point $x_0 \in I$ there exists a straight line L having $L(x_0) = \phi(x_0)$ and lying wholly on or below the graph of ϕ.

L will be called a line of support of ϕ at $x = x_0$. A function ϕ will be called concave on I if $-\phi$ is convex on I. We note that we use the concept of convexity in the weak sense, thus admitting linearity. Furthermore we remark that the above definitions ensure continuity of ϕ on I, except perhaps at its endpoints, if these exist. The less restrictive definition of convexity

$$\phi(\tfrac{1}{2}x_1 + \tfrac{1}{2}x_2) \leqq \tfrac{1}{2}\phi(x_1) + \tfrac{1}{2}\phi(x_2)$$

for all $x_1, x_2 \in I$, does not have this property and on the basis of the axiom of well-ordering it is indeed possible to construct discontinuous functions satisfying the latter definition (cf. [13], 96).

We shall also be concerned with antisymmetrical, concave-convex functions ϕ on I. A function ϕ on I will be called antisymmetrical on I about x_0 if

$$\phi(x_0 + x) + \phi(x_0 - x) = 2\phi(x_0)$$

for some $x_0 \in I$ and all x with $x_0 - x \in I$ and $x_0 + x \in I$; x_0 will be called a central point of ϕ. We note that the property of antisymmetry is confined to the interval I, as is proper. If, for instance, I is closed then any function defined on I is trivially antisymmetric on I about both endpoints of I. An antisymmetrical function ϕ on I will be said to be concave-convex on I if ϕ is concave for $x \leqq x_0$, $x \in I$, and convex for $x \geqq x_0$, $x \in I$, where x_0 is a central point of ϕ. It is clear that such a function will be continuous on I, with the possible exception of the point x_0 and both endpoints of I. If, in addition, we suppose ϕ to be non-decreasing on I then ϕ is also continuous at $x = x_0$.

Finally, we shall say that the distribution given by F is symmetrical about x_0 if

$$F(x_0 + x) + F(x_0 - x) = 1$$

for some x_0 and all real x. This implies that the support I of F is either $(-\infty, +\infty)$ or a finite open or closed interval (a,b) or [a,b] with $x_0 = \tfrac{1}{2}(a + b)$.

2.2. CONVEX TRANSFORMS

The basic result on convex functions is the celebrated JENSEN inequality [19] of which we shall make frequent use in the sequel:

LEMMA 2.2.1

If ϕ is convex on I, then

$$\phi(E\ \underline{x}) \leq E\ \phi(\underline{x}) \quad ,$$

provided both expectations exist. There is equality if and only if ϕ is linear on I.

PROOF

As $\underline{x}$ is non-degenerate and E $\underline{x}$ is finite, E $\underline{x}$ is an interior point of I. Let L be a line of support of ϕ at $x = E\ \underline{x}$. Since $L(x) \leq \phi(x)$ on I and L is linear we have

$$E\ \phi(\underline{x}) \geq E\ L(\underline{x}) = L(E\ \underline{x}) = \phi(E\ \underline{x}) \quad .$$

If there is equality, then $\phi(x) = L(x)$ on a subset of I of probability 1. However, since ϕ is convex on I this means that $\phi(x) = L(x)$ on the smallest interval containing this subset, which is I by definition. The converse is trivial.

If, in addition to being convex, ϕ is also non-decreasing on I a result of a completely different kind may be obtained. Apart from an overall linear change of scale, such a non-decreasing, convex transformation of a random variable effects a contraction of the lower part of the scale of measurement and an extension of the upper part. As, moreover, this deformation increases towards both ends of the scale, the transformation should produce what one intuitively feels to be an increased skewness to the right. Theorem 2.2.1 shows that the best known measures of skewness, being the standardized odd central moments, do indeed react as expected.

THEOREM 2.2.1

If ϕ is a non-decreasing, convex function on I, which is not constant on I, and if $\mu_{2k+1}(\underline{x})$ and $\mu_{2k+1}(\phi(\underline{x}))$ exist, then

$$\frac{\mu_{2k+1}(\underline{x})}{\sigma^{2k+1}(\underline{x})} \leqq \frac{\mu_{2k+1}(\phi(\underline{x}))}{\sigma^{2k+1}(\phi(\underline{x}))} \quad , \quad \text{for } k=1,2,\ldots \quad .$$

PROOF

We start by remarking that ϕ cannot be constant on a set of probability 1, since it would then be constant on I by its monotonicity. As $\underline{x}$ is non-degenerate the variances of $\underline{x}$ and $\phi(\underline{x})$ are positive. Hence without loss of generality we may set $E\,\underline{x} = E\,\phi(\underline{x}) = 0$, $E\,\underline{x}^2 = E\,\phi^2(\underline{x}) \neq 0$, and prove $E\,\underline{x}^{2k+1} \leqq E\,\phi^{2k+1}(\underline{x})$. We drop the trivial case that $\phi(\underline{x}) = \underline{x}$ with probability 1.

Next we collect some facts about the geometry of the situation. According to lemma 2.2.1, since ϕ is not linear on I

$$\phi(0) = \phi(E\,\underline{x}) < E\,\phi(\underline{x}) = 0 \quad .$$

Furthermore $\phi(x) - x$ cannot be non-negative or non-positive for all $x \in I$ since $P(\phi(\underline{x}) \neq \underline{x}) > 0$ and $E\,\phi(\underline{x}) = E\,\underline{x}$. Suppose that e.g. $\phi(x) \leqq x$ for $x < x_0$, and $\phi(x) \geqq x$ for $x \geqq x_0$ for some $x_0 \in I$. Then, as $\phi(x) + x$ is strictly increasing on I and $P(\phi(\underline{x}) \neq \underline{x}) > 0$

$$\int_I (\phi(x) + x)(\phi(x) - x)\, dF(x) \quad > \quad 2x_0 \int_I (\phi(x) - x)\, dF(x) \; = \; 0 \; ,$$

or $E\,\underline{x}^2 < E\,\phi^2(\underline{x})$ which contradicts the standardization of the random variables. In the same way $\phi(x_0) = x_0$, $\phi(x) \geqq x$ for $x \leqq x_0$ and $\phi(x) \leqq x$ for $x \geqq x_0$ leads to the contradiction $E\,\phi^2(\underline{x}) < E\,\underline{x}^2$. It follows that $\phi(x) - x$ has at least two changes of sign on I; since it is convex it changes sign exactly twice, say for $x_1 < x_2 \in I$.

Let us suppose first that ϕ is continuous on I. Because $\phi(0) < 0$ and ϕ is convex on I we find that $x_1 < 0 < x_2$, $\phi(x_1) = x_1$,

$\phi(x_2) = x_2$, $\phi(x) < x$ for $x_1 < x < x_2$, and $\phi(x) > x$ for $x < x_1$ and $x > x_2$. Thus we have the geometrical situation sketched in figure 2.2.1 where x_3, $0 < x_3 < x_2$ denotes the unique point where $\phi(x) = 0$.

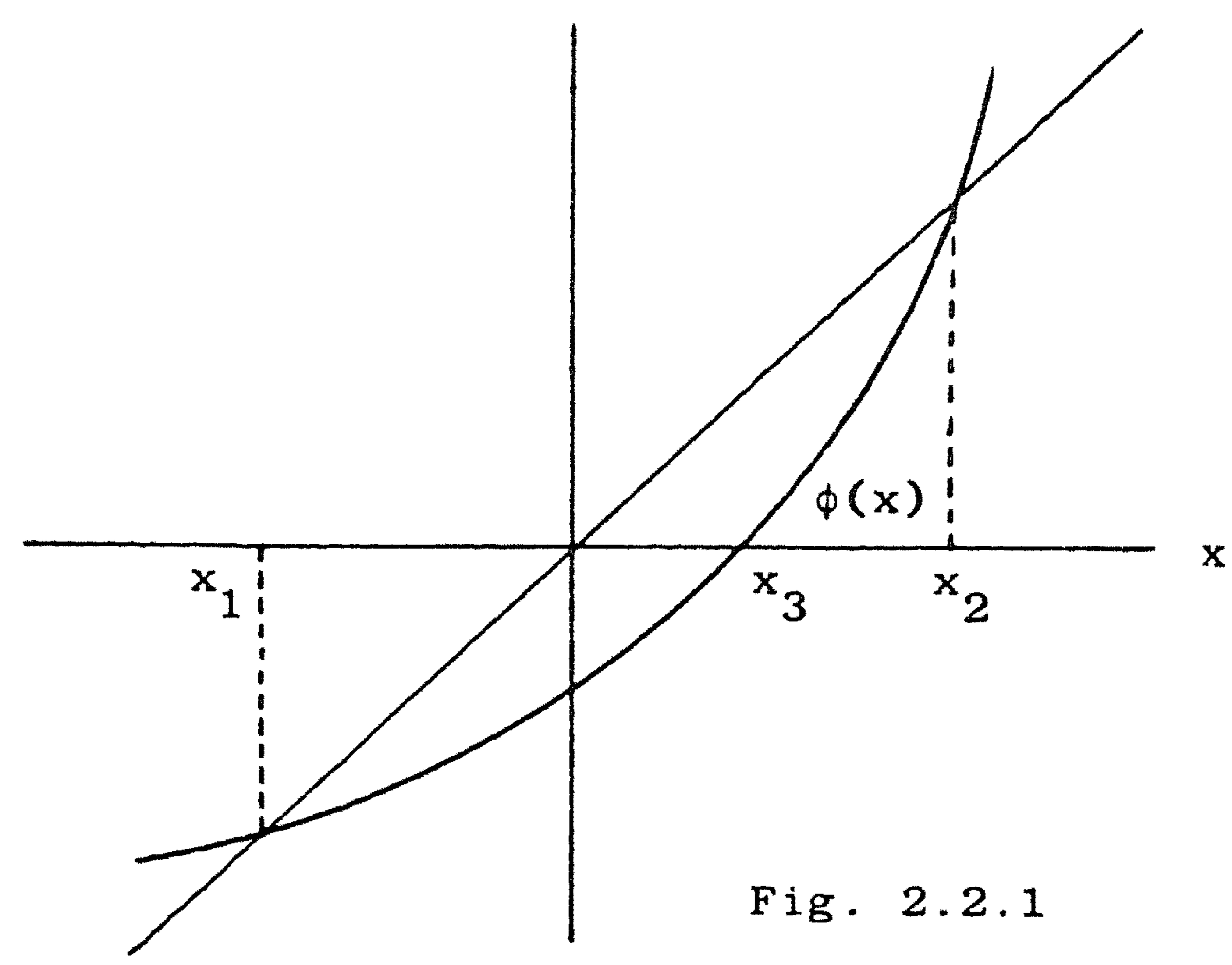


Fig. 2.2.1

Consider the functions

$$(2.2.1)\qquad \psi_k(x) = \frac{\phi^{2k+1}(x) - x^{2k+1}}{\phi(x) - x} = \sum_{i=0}^{2k} \phi^i(x)\, x^{2k-i} =$$

$$= (\phi(x) + x) \sum_{i=0}^{k-1} \phi^{2i}(x)\, x^{2k-2i-1} + \phi^{2k}(x) =$$

$$= (\phi(x) + x) \sum_{i=1}^{k} \phi^{2i-1}(x)\, x^{2k-2i} + x^{2k}\ , \quad \text{and}$$

$$(2.2.2)\qquad \chi_k(x) = f_k(x;x_1,x_2) = \frac{2k+1}{2}\, \frac{x_2^{2k} - x_1^{2k}}{x_2 - x_1} (\phi(x) + x) +$$

$$+ (2k+1) \frac{x_2 x_1^{2k} - x_1 x_2^{2k}}{x_2 - x_1}\ ,$$

where the second member of (2.2.1) is defined for $x = x_1$ and $x = x_2$ by continuity.

Clearly $\psi_k(x) > 0$ on I, and

$$\chi_k(x_1) = \psi_k(x_1) = (2k+1)x_1^{2k}$$

$$\chi_k(x_2) = \psi_k(x_2) = (2k+1)x_2^{2k} \quad .$$

χ_k may thus be called a $(\phi(x) + x)$- chord of ψ_k (cf. [13], 75). We shall show that $\psi_k(x) \leqq \chi_k(x)$ for $x_1 \leqq x \leqq x_2$ and $\psi_k(x) \geqq \chi_k(x)$ for $x \leqq x_1$ and $x \geqq x_2$.

First we consider $f_k(x;\xi_1,\xi_2)$ for fixed $x \in I$ as a function of ξ_1 and ξ_2 in the rectangle $x_1 \leqq \xi_1 \leqq 0$, $0 \leqq \xi_2 \leqq x_2$. Differentiating with respect to ξ_1 and ξ_2 respectively we obtain

$$\frac{\partial f_k(x;\xi_1,\xi_2)}{\partial \xi_1} = \frac{(2k+1)((2k-1)\xi_1^{2k} - 2k\xi_1^{2k-1}\xi_2 + \xi_2^{2k})}{2(\xi_2 - \xi_1)^2} (\phi(x) + x - 2\xi_2)$$

$$\gtreqless 0 \quad \text{for } \phi(x) + x - 2\xi_2 \gtreqless 0, \quad \text{and}$$

$$\frac{\partial f_k(x;\xi_1,\xi_2)}{\partial \xi_2} = \frac{(2k+1)((2k-1)\xi_2^{2k} - 2k\xi_2^{2k-1}\xi_1 + \xi_1^{2k})}{2(\xi_2 - \xi_1)^2} (\phi(x) + x - 2\xi_1)$$

$$\gtreqless 0 \quad \text{for } \phi(x) + x - 2\xi_1 \gtreqless 0 \quad .$$

We may now compare $\chi_k(x)$ and $\psi_k(x)$ for all $x \in I$. We distinguish five sub-intervals of I:

a) $x \geqq x_2$, $x \in I$.

We set $\xi_2 = x_2$. As $\phi(x) + x - 2x_2 \geqq 0$, $f_k(x;\xi_1,x_2)$ is a non-decreasing function of ξ_1 for $x_1 \leqq \xi_1 \leqq 0$, hence

$$\chi_k(x) = f_k(x;x_1,x_2) \leqq f_k(x;0,x_2) = \tfrac{1}{2}(2k+1)x_2^{2k-1}(\phi(x) + x) \leqq$$

$$\leqq (\phi(x) + x) \sum_{i=0}^{k-1} \phi^{2i}(x)\, x^{2k-2i-1} + \phi^{2k}(x) = \psi_k(x) \quad ,$$

since $0 < x_2 \leqq x \leqq \phi(x)$.

b) $x \leqq x_1$, $x \in I$.

We set $\xi_1 = x_1$. As $\phi(x) + x - 2x_1 \leqq 0$, $f_k(x;x_1,\xi_2)$ is a non-increasing function of ξ_2 for $0 \leqq \xi_2 \leqq x_2$, hence

$$\chi_k(x) = f_k(x;x_1,x_2) \leqq f_k(x;x_1,0) = \tfrac{1}{2}(2k+1)x_1^{2k-1}(\phi(x) + x) \leqq$$

$$\leqq (\phi(x) + x) \sum_{i=1}^{k} \phi^{2i-1}(x)\, x^{2k-2i} + x^{2k} = \psi_k(x) \quad ,$$

since $x \leqq \phi(x) \leqq x_1 < 0$.

c) $x_3 \leqq x \leqq x_2$.

We set $\xi_2 = x_2$. As $\phi(x) + x - 2x_2 \leqq 0$, $f_k(x;\xi_1,x_2)$ is a non-increasing function of ξ_1 for $x_1 \leqq \xi_1 \leqq 0$, hence

$$\chi_k(x) = f_k(x;x_1,x_2) \geqq f_k(x;0,x_2) = \tfrac{1}{2}(2k+1)x_2^{2k-1}(\phi(x) + x) \geqq$$

$$\geqq (\phi(x) + x) \sum_{i=0}^{k-1} \phi^{2i}(x)\, x^{2k-2i-1} + \phi^{2k}(x) = \psi_k(x) \quad ,$$

since $0 \leqq \phi(x) \leqq x \leqq x_2$.

d) $x_1 \leqq x \leqq 0$.

We set $\xi_1 = x_1$. As $\phi(x) + x - 2x_1 \geqq 0$, $f_k(x;x_1,\xi_2)$ is a non-decreasing function of ξ_2 for $0 \leqq \xi_2 \leqq x_2$, hence

$$\chi_k(x) = f_k(x;x_1,x_2) \geqq f_k(x;x_1,0) = \tfrac{1}{2}(2k+1)x_1^{2k-1}(\phi(x) + x) \geqq$$

$$\geqq (\phi(x) + x) \sum_{i=1}^{k} \phi^{2i-1}(x)\, x^{2k-2i} + x^{2k} = \psi_k(x) \quad ,$$

since $x_1 \leqq \phi(x) \leqq x \leqq 0$.

e) $0 \leq x \leq x_3$.

We consider a fixed value of x and let (ξ_1, ξ_2) range through the rectangle

$$x_1 \leq \xi_1 \leq \phi(x) \leq 0 \quad \text{and} \quad 0 \leq x \leq \xi_2 \leq x_2 .$$

For all (ξ_1, ξ_2) in this set

$$\phi(x) + x - 2\xi_2 \leq 2x - 2\xi_2 \leq 0 \quad \text{and} \quad \phi(x) + x - 2\xi_1 \geq 2\phi(x) - 2\xi_1 \geq 0.$$

Hence on this set $f_k(x;\xi_1,\xi_2)$ is a non-increasing function of ξ_1 and a non-decreasing function of ξ_2, and therefore

$$\chi_k(x) = f_k(x;x_1,x_2) \geq f_k(x;\phi(x),x) =$$

$$= \frac{2k+1}{2} \frac{x^{2k+1} - \phi^{2k+1}(x) + x\phi^{2k}(x) - \phi(x)x^{2k}}{x - \phi(x)} \geq$$

$$\geq \frac{2k+1}{2} \frac{\phi^{2k+1}(x) - x^{2k+1}}{\phi(x) - x} = \psi_k(x) ,$$

since $\phi(x) \leq 0 \leq x$, $k \geq 1$ and $\psi_k(x) > 0$.

So far we have found that, if ϕ is continuous on I, then for all $x \in I$ the functions $\psi_k(x) - \chi_k(x)$ and $\phi(x) - x$ have the same sign so their product is non-negative on I. Hence

$$\int_I (\psi_k(x) - \chi_k(x)) (\phi(x) - x) \, dF(x) \geq 0 ,$$

and as a result

$$E\,\phi^{2k+1}(\underline{x}) - E\,\underline{x}^{2k+1} = \int_I (\phi^{2k+1}(x) - x^{2k+1}) \, dF(x) =$$

$$= \int_I \psi_k(x) (\phi(x) - x) \, dF(x) \geq \int_I \chi_k(x) (\phi(x) - x) \, dF(x) =$$

$$= \frac{2k+1}{2} \; \frac{x_2^{2k} - x_1^{2k}}{x_2 - x_1} \int_I (\phi^2(x) - x^2)\, dF(x) + (2k+1) \frac{x_2 x_1^{2k} - x_1 x_2^{2k}}{x_2 - x_1} \,\cdot$$

$$\cdot \int_I (\phi(x) - x)\, dF(x) = 0 \,,$$

since $E\ \phi^2(\underline{x}) = E\ \underline{x}^2$ and $E\ \phi(\underline{x}) = E\ \underline{x}$.

The case that ϕ is not continuous on I remains to be considered. As ϕ is non-decreasing and convex on I it can only have a discontinuity at the right endpoint of I, if this exists. If x_2 is an interior point of I the proof remains unchanged. If, however, x_2 should be the right endpoint of I, then $\phi(x) < x$ for $0 \leqq x < x_2$, $\phi(x_2) > x_2$, and $\psi_k(x_2) \neq \chi_k(x_2)$ accordingly. Furthermore, the reasoning given under c) does not continue to hold for $x = x_2$, but from a) we have $\psi_k(x_2) > \chi_k(x_2)$. The remainder of the proof remains unchanged. This completes the proof of theorem 2.2.1.

Going over the proof given we note that we have extensively exploited the fact that $E\ \phi(\underline{x}) = E\ \underline{x}$ and $E\ \phi^2(\underline{x}) = E\ \underline{x}^2$. We have not only made use of these properties to arrive at the geometry of the situation but the crucial part of the proof depends entirely on these points. In contrast to this, we have hardly made use of the fact that $E\ \underline{x} = 0$. We have in fact only needed this to show that $\phi(0) \leqq 0$. It should therefore be possible to extend the result of the theorem somewhat so as to cover moments that are not centered at the expectation under suitable assumptions. We shall not, however, go into this any further.

A continuous version of theorem 2.2.1 may be obtained as follows. If ϕ is non-decreasing, convex on I then it is easy to show that for $0 \leqq \lambda \leqq \lambda' \leqq 1$, $\lambda'\phi(x) + (1-\lambda')x$ is a non-decreasing, convex function of $\lambda\phi(x) + (1-\lambda)x$ on I. As a result, the statement that

$$\frac{\mu_{2k+1}(\lambda\phi(\underline{x}) + (1-\lambda)\underline{x})}{\sigma^{2k+1}(\lambda\phi(\underline{x}) + (1-\lambda)\underline{x})}$$

is a non-decreasing function of λ for $0 \leq \lambda \leq 1$ is equivalent to the inequality of theorem 2.2.1.

It may be appropriate at this stage to remark that if $\phi(x)$ is a non-decreasing, convex function on I, then $\phi(-x)$, $-\phi(-x)$ and $-\phi(x)$ are non-increasing convex, non-decreasing concave and non-increasing concave functions of x. Hence, to every theorem on convex or non-decreasing convex functions there correspond analogous theorems on concave or non-increasing convex, non-decreasing concave and non-increasing concave functions. In lemma 2.2.1 and theorem 2.2.1 for instance, the inequalities are simply reversed if we replace convex by concave. In what follows we shall not, in general, formulate these dual results explicitly. The same situation does, of course, exists when we treat antisymmetrical, concave-convex transforms in the sequel.

From this remark it is also obvious that no general statement like theorem 2.2.1 can be made about the even moments of a non-decreasing, convex transform $\phi(\underline{x})$. If, for instance, for all non-decreasing, convex ϕ, $E\,\underline{x}^{2k} \leq E\,\phi^{2k}(\underline{x})$ would hold after standardization, then also $E(-\phi(\underline{x}))^{2k} \geq E(-\underline{x})^{2k}$. However, if φ happens to be continuous and strictly increasing, $-x$ is an increasing, convex function of $-\phi(x)$, and we have a contradiction unless both expectations are equal.

From the reasoning given before the statement of theorem 2.2.1 one might expect that a similar theorem would hold for any measure of skewness. Surprisingly enough this is not the case as may be seen from the following counter-example.

K. PEARSON appears to have introduced the measure of skewness

$$s(\underline{x}) = \frac{E\,\underline{x} - m(\underline{x})}{\sigma(\underline{x})} ,$$

which was studied subsequently by H. HOTELLING and L.M. SOLOMONS [18] and R. CARVER [6] who showed that $|s(\underline{x})| \leq 1$ for any distribution. We shall demonstrate that a distribution F of $\underline{x}$ exists for which

$$s(\underline{x}) > s(\varphi(\underline{x}))$$

for every non-decreasing, convex function ϕ which is not linear on I.

Let the random variable $\underline{x}$ assume the values $x_1 < x_2 < x_3$ with positive probabilities p_1, p_2, p_3, $p_1 + p_2 + p_3 = 1$, and let $p_3 > \frac{1}{2}$. Let ϕ be non-decreasing, convex and non-linear on $I = [x_1, x_3]$ and let $\underline{x}$ and $\phi(\underline{x})$ be standardized in such a way that

$$E\,\underline{x} = E\,\phi(\underline{x}) = 0 \;, \quad E\,\underline{x}^2 = E\,\phi^2(\underline{x}) \;.$$

From the proof of theorem 2.2.1 we know that $\phi(x) - x$ changes sign twice on I and that $\phi(x_3) > x_3$. Since $m(\underline{x}) = x_3$ and $m(\phi(\underline{x})) = \phi(x_3)$ we have $s(\phi(\underline{x})) < s(\underline{x})$.

The present author is inclined to conclude that $s(\underline{x})$ is not a very satisfactory measure of skewness for distributions of this type.

2.3. ANTISYMMETRICAL CONCAVE-CONVEX TRANSFORMS

For antisymmetrical, concave-convex functions ϕ on I we shall first prove a theorem analogous to lemma 2.2.1. The gist of this result is simply that if we impose suitable restrictions on the distribution F of $\underline{x}$, we can make sure that the convex part of ϕ plays a dominant role and hence that JENSEN's inequality continues to hold.

THEOREM 2.3.1

Let ϕ be an antisymmetrical, concave-convex function on I and let x_0 be a central point of ϕ. If, for $x \geqq 0$, $F(x_0+x) + F(x_0-x)$ is a non-decreasing function of x, then

$$\phi(E\,\underline{x}) \;\leqq\; E\,\phi(\underline{x}) \quad ,$$

provided both expectations exist. There is equality if and only if for some $c \geqq 0$, $\phi(x_0+x)$ is linear for $|x| \leqq c$ (or $|x| < c$) and $F(x_0+x) + F(x_0-x) = 1$ for $|x| > c$ (or $|x| \geqq c$ respectively).

PROOF

Without loss of generality we set $\phi(x_0) = x_0 = 0$. We note that, since $H(x) = F(x) + F(-x)$ is non-decreasing for $x \geq 0$,

$$E\,\underline{x} = \int_I x\,dF(x) = \int_0^\infty x\,dH(x) \geq 0 \quad . \tag{2.3.1}$$

Furthermore we have for $x \geq 0$, and therefore for all x,

$$H(x) \leq \lim_{x\to\infty} H(x) = 1 \quad . \tag{2.3.2}$$

Hence, if $x \leq 0$, $x \in I$, then $F(x) > 0$ and $F(-x) < 1$, or $-x \in I$.

Let us suppose first that $E\,\underline{x} = 0$. By (2.3.1) and (2.3.2) we have $H(x) = 1$ for $x > 0$, and hence $H(x) = 1$ for all x, i.e. the distribution given by F is symmetrical about $x = 0$. Therefore $\phi(E\,\underline{x}) = 0 = E\,\phi(\underline{x})$ which proves the theorem together with the condition for equality for the case that $E\,\underline{x} = 0$.

For the remainder of the proof we may therefore suppose that $E\,\underline{x} > 0$. Let L be a line of support of ϕ for $x \geq 0$ at $x = E\,\underline{x}$. Then

$$\phi(x) - L(x) \geq 0 \qquad \text{for } x \geq 0,\ x \in I, \text{ and}$$

$$L(x) + L(-x) = 2L(0) \leq 0 = \phi(x) + \phi(-x)\ , \qquad \text{or}$$

$$\phi(x) - L(x) \geq L(-x) - \phi(-x) \qquad \text{for } x \leq 0,\ x \in I.$$

Hence

$$\int_I (\phi(x) - L(x))\,dF(x) \geq \tag{2.3.3}$$

$$\geq \int_0^\infty (\phi(x) - L(x))\,dF(x) + \int_{-\infty}^0 (L(-x) - \phi(-x))\,dF(x) =$$

$$= \int_0^\infty (\phi(x) - L(x))\,dH(x) \geq 0\ , \qquad \text{or}$$

$$E\,\phi(\underline{x}) \geq E\,L(\underline{x}) = L(E\,\underline{x}) = \phi(E\,\underline{x}) \quad .$$

If ϕ is linear for $|x| \leqq c$ and $H(x) = 1$ for $|x| > c$, it is obvious that

$$0 < E\,\underline{x} = \int_0^{\infty} x \, dH(x) = \int_0^{c+} x \, dH(x) \leqq c \ .$$

Hence L may be chosen in such a way that $L(x) = \phi(x)$ for $|x| \leqq c$. Then

$$\phi(x) - L(x) = L(-x) - \phi(-x) \qquad \text{for } x \leqq 0,\ x \in I \ . \tag{2.3.4}$$

Since for $|x| > c$, $x \in I$ if and only if $-x \in I$ by the assumption $H(x) = 1$ for $|x| > c$, we have

$$\int_I (\phi(x) - L(x)) \, dF(x) =$$

$$= \int_{c+}^{\infty} (\phi(x) - L(x)) \, dF(x) + \int_{-\infty}^{-c-} (L(-x) - \phi(-x)) \, dF(x) =$$

$$= \int_{c+}^{\infty} (\phi(x) - L(x)) \, dH(x) = 0 \ .$$

The same method of proof may be used if ϕ is linear for $|x| < c$ and $H(x) = 1$ for $|x| \geqq c$.

Conversely, if $\phi(E\,\underline{x}) = E\,\phi(\underline{x})$ then necessarily $L(0) = 0$, since otherwise the first inequality in (2.3.3) would be strict because $0 = x_0 \in I$ (cf. section 2.1) and hence $F(0) > 0$. This implies that (2.3.4) holds and that $\phi(x) = L(x)$ for $|x| \leqq E\,\underline{x}$. Hence for some real number $c \geqq E\,\underline{x} > 0$, $\phi(x) = L(x)$ for $|x| \leqq c$ (or $|x| < c$) and $\phi(x) > L(x)$ for $x > c$ (or $x \geqq c$ respectively). From (2.3.1) and (2.3.2) we then have $H(x) = 1$ for $|x| > c$ (or $|x| \geqq c$ respectively). This completes the proof of theorem 2.3.1.

We note that the conditions for equality in theorem 2.3.1 are already much more involved than those of lemma 2.2.1. However, two extreme cases where the conditions are satisfied are obvious. For $c = \infty$ they reduce to linearity of ϕ on I, whereas for $c = 0$ it is required that the distribution of $\underline{x}$ be symmetric about $x = x_0$. The case that $\phi(x_0+x)$ is linear for $|x| < c$ (and not for $|x| \leqq c$) and $F(x_0+x) + F(x_0-x) = 1$ for $|x| \geqq c$ can only occur if either ϕ is linear on I and I is open, or if $I = [x_0-c, x_0+c]$, ϕ is discontinuous for $x = x_0+c$ and $x = x_0-c$, and $P(\underline{x} = x_0+c) = P(\underline{x} = x_0-c) > 0$.

If we consider non-decreasing, antisymmetrical, concave-convex transforms of a symmetrically distributed random variable we may prove a result in the same spirit as theorem 2.2.1. Roughly speaking, such a transformation carries probability mass to the tails of the distribution and consequently the following theorem on the standardized even central moments of $\underline{x}$ and $\phi(\underline{x})$ is intuitively obvious.

THEOREM 2.3.2

Let ϕ be a non-decreasing, antisymmetrical, concave-convex function on I, which is not constant on I, and let the distribution given by F be symmetrical about x_0, where x_0 denotes a central point of ϕ. Then, if $E\ \phi^{2k}(\underline{x})$ exists,

$$\frac{\mu_{2k}(\underline{x})}{\sigma^{2k}(\underline{x})} \leqq \frac{\mu_{2k}(\phi(\underline{x}))}{\sigma^{2k}(\phi(\underline{x}))}, \qquad \text{for } k=2,3,\ldots\ .$$

PROOF

The existence of $E\ \phi^{2k}(\underline{x})$ clearly implies existence of $E\ \underline{x}^{2k}$; furthermore, ϕ is not constant on a set of probability 1 because of its monotonicity, and since $\underline{x}$ is non-degenerate $\underline{x}$ and $\phi(\underline{x})$ have finite, positive variances. Hence without loss of generality we may set $x_0 = \phi(x_0) = 0$, or $E\ \underline{x} = E\ \phi(\underline{x}) = 0$, and $E\ \underline{x}^2 = E\ \phi^2(\underline{x}) \neq 0$, and proceed to prove that $E\ \underline{x}^{2k} \leqq E\ \phi^{2k}(\underline{x})$. We disregard the trivial case that $\phi(\underline{x}) = \underline{x}$ with probability 1.

Now $\phi(x) - x$ cannot be non-negative or non-positive for all $x \geqq 0$, $x \in I$, for in that case $\phi^2(x) - x^2$ would be non-negative or non-positive for all $x \in I$; since we have supposed that $P(\phi(\underline{x}) \neq \underline{x}) > 0$,

this would mean that $E\,\phi^2(\underline{x}) - E\,\underline{x}^2$ would not be equal to zero. As ϕ is convex for $x \geqq 0$, $x \in I$, and $\phi(0) = 0$, it follows that $\phi(x) - x \leqq 0$ for $0 \leqq x < x_0'$ and $\phi(x) - x \geqq 0$ for $x \geqq x_0'$, $x \in I$, for some $x_0' > 0$, $x_0' \in I$. Hence $\phi^2(x) - x^2 \leqq 0$ for $|x| < x_0'$ and $\phi^2(x) - x^2 \geqq 0$ for $|x| \geqq x_0'$, $x \in I$.
Writing

$$\phi^{2k}(x) - x^{2k} = \psi_k(x)\ (\phi^2(x) - x^2)$$

and noting that

$$\psi_k(x) = \sum_{j=0}^{k-1} \phi^{2j}(x)\ x^{2k-2j-2}$$

is a non-negative, even function on I which is increasing for $x \geqq 0$, we find

$$E\,\phi^{2k}(\underline{x}) - E\,\underline{x}^{2k} = \int_I \psi_k(x)\ (\phi^2(x) - x^2)\ dF(x) \geqq$$

$$\geqq \psi_k(x_0') \int_I (\phi^2(x) - x^2)\ dF(x) = 0 \quad ,$$

which completes the proof.

Analogous to theorem 2.2.1 we may formulate a continuous version of theorem 2.3.2 stating that under the conditions of the theorem

$$\frac{\mu_{2k}(\lambda\phi(\underline{x}) + (1-\lambda)\underline{x})}{\sigma^{2k}(\lambda\phi(\underline{x}) + (1-\lambda)\underline{x})}$$

is a non-decreasing function of λ for $0 \leq \lambda \leq 1$. One may, however, also generalize theorem 2.3.2 in another direction. For real $a \geqq 0$ let $\nu_a(\underline{x})$ denote the a-th absolute central moment

$$\nu_a(\underline{x}) = E\,|\underline{x} - E\,\underline{x}|^a \ .$$

Then under the conditions of the theorem

$$\frac{\nu_a^b(\underline{x})}{\nu_b^a(\underline{x})} \leqq \frac{\nu_a^b(\phi(\underline{x}))}{\nu_b^a(\phi(\underline{x}))} \qquad \text{for } 0 \leqq b \leqq a \ ,$$

provided $\nu_a(\phi(\underline{x}))$ exists. Standardizing in such a way that $E\ \underline{x} = E\ \phi(\underline{x}) = 0$ and $\nu_b(\underline{x}) = \nu_b(\phi(\underline{x}))$ the proof follows the same pattern as the proof of the theorem, since one may show by the mean value theorem that for $x \geqq 0$

$$\psi(x) = \frac{\phi^a(x) - x^a}{\phi^b(x) - x^b} = \frac{a}{b}(\theta x^b + (1-\theta)\phi^b(x))^{\frac{a-b}{b}} \ \substack{\geq \\ = \\ <} \ \frac{a}{b} x_0'^{\,a-b} \qquad \text{for } x \ \substack{\geq \\ = \\ <} \ x_0' \ .$$

We note in passing that a similar device may be used to prove theorem 2.2.1. In that case the inequality

$$(\psi_k(x) - \chi_k(x))(\phi(x) - x) \ \geqq \ 0 \qquad \text{on I}$$

may be obtained by application of the mean value theorem to

$$\frac{(\phi^{2k+1}(x) - A\ \phi^2(x)) - (x^{2k+1} - A\ x^2)}{\phi(x) - x} \ ,$$

where

$$A = \frac{2k+1}{2} \frac{x_2^{2k} - x_1^{2k}}{x_2 - x_1} \ .$$

Chapter 3

ORDER STATISTICS

3.1. NOTATION

Suppose that F is continuous on I and let $\underline{x}_{1:n} < \underline{x}_{2:n} < \ldots < \underline{x}_{n:n}$ denote an ordered sample of size n from the distribution F; $\underline{x}_{i:n}$ is called the i-th order statistic of a sample of size n from F. The distribution function $F_{i:n}$ of $\underline{x}_{i:n}$ is given by (cf. [23], 12)

$$(3.1.1)\qquad F_{i:n}(x) = \frac{n!}{(i-1)!(n-i)!}\int_0^{F(x)} y^{i-1}(1-y)^{n-i}\,dy = B_{i:n}F(x),^{\dagger)}$$

where $B_{i:n}$ denotes the incomplete beta function

$$(3.1.2)\qquad B_{i:n}(y) = \frac{n!}{(i-1)!(n-i)!}\int_0^{y} u^{i-1}(1-u)^{n-i}\,du \qquad \text{for integer } 1 \leqq i \leqq n\ .$$

The inverse function G of F is defined for $0 < y < 1$ except for discontinuities by

$$(3.1.3)\qquad GF(x) = x \qquad \text{for } x \in I\ .$$

Hence for any function ψ on I for which $E\,\psi(\underline{x}_{i:n})$ exists we have

---

†) We shall usually not use brackets to denote composite functions like this, and write e.g. $B_{i:n}F$ and $B_{i:n}F(x)$ rather than $B_{i:n}(F(.))$ and $B_{i:n}(F(x))$.

$$(3.1.4)\quad E\,\psi(\underline{x}_{i:n}) = \frac{n!}{(i-1)!(n-i)!}\int_I \psi(x)\,F^{i-1}(x)\,(1-F(x))^{n-i}\,dF(x) =$$

$$= \int_0^1 \psi G(y)\; b_{i:n}(y)\,dy \quad ,$$

where

$$(3.1.5)\qquad b_{i:n}(y) = B'_{i:n}(y) = \frac{n!}{(i-1)!(n-i)!}\,y^{i-1}\,(1-y)^{n-i} \quad .$$

In the remaining part of this study we shall confine our attention throughout to the class $\mathcal{F}$ of distribution functions F satisfying

(3.1.6) F is twice differentiable on I with continuous second derivative F'' on I,

(3.1.7) $F'(x) > 0$ on I,

(3.1.8) There exist integers i and n, $1 \leqq i \leqq n$, such that $E\,\underline{x}_{i:n}$ exists.

These conditions imply that the inverse function G of F is uniquely defined for $0 < y < 1$ by (3.1.3) and

(3.1.9) G is twice differentiable on (0,1) with continuous second derivative G'' on (0,1),

(3.1.10) $G'(y) > 0$ for $0 < y < 1$,

(3.1.11) There exist non-negative integers a and b such that $|G(y)\,y^a\,(1-y)^b|$ is bounded for $0 < y < 1$.

It is easy to see that, conversely, (3.1.9) - (3.1.11) imply (3.1.6) - (3.1.8) and that therefore (3.1.9) - (3.1.11) may also be used to define $\mathcal{F}$.

In this chapter we shall repeatedly meet the central moments $\mu_j(i,n)$, the absolute central moments $\nu_j(i,n)$ and the median $m(i,n)$ of the beta distribution $B_{i:n}$. We define these by (cf. section 2.1)

$$(3.1.12) \qquad \mu_j(i,n) = \int_0^1 (y - \frac{i}{n+1})^j \, b_{i:n}(y)\, dy \quad ,$$

$$(3.1.13) \qquad \nu_j(i,n) = \int_0^1 |y - \frac{i}{n+1}|^j \, b_{i:n}(y)\, dy \quad , \text{ and}$$

$$(3.1.14) \qquad B_{i:n}(m(i,n)) = \frac{1}{2} \quad .$$

3.2. LARGE SAMPLE PROPERTIES

In this section we establish asymptotic expressions for $E\, \underline{x}_{i:n}$, $F(E\, \underline{x}_{i:n})$, $\mu_k(\underline{x}_{i:n})$ and $m(\underline{x}_{i:n})$ as i and n tend to infinity in such a way that $\lim \frac{i}{n} = r$, $0 < r < 1$.

The result for $E\, \underline{x}_{i:n}$ is well known (cf. [10] and [20]); the result for $F(E\, \underline{x}_{i:n})$, which is derived from it, closely resembles the corresponding expression given by G. BLOM in [4]. The expression for $\mu_k(\underline{x}_{i:n})$ for odd values of k seems to be new in this generality, although the special case $k = 3$ has been considered in [10]. All of these results are obtained by expanding $G(y)$ in (3.1.4) about $y = \frac{i}{n+1}$.

The main reason that we give the derivations of these results in full detail is that most of the previous proofs of similar expressions lack rigour as G. BLOM has noted ([4], 48). BLOM's own results for $E\, \underline{x}_{i:n}$ and $F(E\, \underline{x}_{i:n})$ are established under slightly different conditions from the ones required here, though the method of proof closely resembles his.

The expression for $m(\underline{x}_{i:n})$ does not seem to have been published previously, although it is connected with known results. The rather surprising simplicity of this expression has a special significance for the approach to the problem of plotting on probability papers described in [2] as will be shown at the end of this section.

We start by establishing asymptotic expressions for the moments and median of the beta distribution.

LEMMA 3.2.1

If, as n tends to infinity, $\lim \frac{i}{n} = r$, $0 < r < 1$, then

$$\mu_j(i,n) = \frac{j!}{(\frac{j}{2})!} 2^{-\frac{j}{2}} \left(\frac{i(n+1-i)}{(n+1)^3}\right)^{\frac{j}{2}} + \mathcal{O}(n^{-\frac{j}{2}-1}) =$$

$$= \mathcal{O}(n^{-\frac{j}{2}}) \quad \text{if } j \text{ is even },$$

$$\mu_j(i,n) = \mathcal{O}(n^{-\frac{j+1}{2}}) \quad \text{if } j \text{ is odd },$$

$$\nu_j(i,n) = \mathcal{O}(n^{-\frac{j}{2}}) \quad , \text{ and}$$

$$m(i,n) = \frac{i-\frac{1}{3}}{n+\frac{1}{3}} + \mathcal{O}(n^{-\frac{3}{2}+\varepsilon}) , \quad \text{for all } \varepsilon > 0 .$$

PROOF

It is well known that under the conditions stated the standardized form of the distribution $B_{i:n}$ tends to the standard normal one (cf. [8], 252). From [21], 184, we find that this implies convergence of the standardized moments of $B_{i:n}$ to those of the standard normal distribution:

$$\lim \frac{\mu_j(i,n)}{(\mu_2(i,n))^{\frac{j}{2}}} = \frac{j!}{(\frac{j}{2})!} 2^{-\frac{j}{2}} \quad , \text{ if } j \text{ is even },$$

$$= 0 \quad , \text{ if } j \text{ is odd} .$$

Since

$$\text{(3.2.1)} \qquad \mu_2(i,n) = \frac{i(n+1-i)}{(n+1)^2(n+2)} = \frac{i(n+1-i)}{(n+1)^3} + \mathcal{O}(n^{-2}) = \mathcal{O}(n^{-1}),$$

and the $\mu_j(i,n)$ are rational functions of i and n the result for $\mu_j(i,n)$ follows. For $\nu_j(i,n)$ the proof is analogous. A direct proof of these results may be given by a method used in [3].

To obtain the result of the lemma for m(i,n) the asymptotic normality of the distribution $B_{i:n}$ is obviously insufficient and one has to establish a second order term. Let the random variable $\underline{y}$ be distributed according to $B_{i:n}$ and consider the random variable

$$\underline{z} = (\underline{y} - \frac{i-1}{n-1})\sqrt{\frac{(n-1)^3}{(i-1)(n-i)}} \quad ;$$

according to the first part of the lemma all central moments of $\underline{z}$ remain bounded as n tends to infinity, and as E $\underline{z}$ tends to zero the same is true for the moments about zero. If by h we denote the probability density of $\underline{z}$ we have

$$h(z) = \sqrt{\frac{(i-1)(n-i)}{(n-1)^3}}\; b_{i:n}\left(z\sqrt{\frac{(i-1)(n-i)}{(n-1)^3}} + \frac{i-1}{n-1}\right) ,$$

or after some rearrangement

$$\begin{aligned}\log h(z) = {} & (i-\tfrac{1}{2}) \log (i-1) + (n-i+\tfrac{1}{2}) \log (n-i) - (n+\tfrac{1}{2}) \log (n-1) + \\ & + \log \frac{n!}{(i-1)!(n-i)!} + (i-1) \log \left\{1 + z\sqrt{\frac{n-i}{(i-1)(n-1)}}\right\} + \\ & + (n-i) \log \left\{1 - z\sqrt{\frac{i-1}{(n-i)(n-1)}}\right\} .\end{aligned}$$

In order to expand the factorials we make use of

$$\log \Gamma(x+a) = (x+a-\tfrac{1}{2}) \log x - x + \tfrac{1}{2} \log 2\pi + \mathcal{O}(x^{-1}) \quad \text{for } x \to \infty \text{ and}$$

fixed a (cf. [25], chapter 13) to obtain

$$\begin{aligned}\log \Gamma(n+1) &= (n+\tfrac{1}{2}) \log (n-1) - (n-1) + \tfrac{1}{2} \log 2\pi + \mathcal{O}(n^{-1}) \\ \log \Gamma(i) &= (i-\tfrac{1}{2}) \log (i-1) - (i-1) + \tfrac{1}{2} \log 2\pi + \mathcal{O}(n^{-1}) \\ \log \Gamma(n-i+1) &= (n-i+\tfrac{1}{2}) \log (n-i) - (n-i) + \tfrac{1}{2} \log 2\pi + \mathcal{O}(n^{-1}) .\end{aligned}$$

Expanding the logarithms we find that uniformly for $|z| \leqq n^{\frac{1}{5}\varepsilon}$, $0 < \varepsilon < \frac{5}{2}$,

$$(i-1) \log \left\{1 + z\sqrt{\frac{n-i}{(i-1)(n-1)}}\right\} + (n-i) \log \left\{1 - z\sqrt{\frac{i-1}{(n-i)(n-1)}}\right\} =$$

$$= -\frac{1}{2}z^2 - \frac{1}{3}z^3 \frac{2i-n-1}{\sqrt{(i-1)(n-i)(n-1)}} + \mathcal{O}(z^4 n^{-1}) =$$

$$= -\frac{1}{2}z^2 - \frac{1}{3}z^3 \frac{2i-n-1}{\sqrt{(i-1)(n-i)(n-1)}} + \mathcal{O}(n^{-1+\frac{4}{5}\varepsilon}) ,$$

hence

$$\log h(z) = -\frac{1}{2} \log 2\pi - \frac{1}{2}z^2 - \frac{1}{3}z^3 \frac{2i-n-1}{\sqrt{(i-1)(n-i)(n-1)}} + \mathcal{O}(n^{-1+\frac{4}{5}\varepsilon}),$$

or

$$h(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \left\{1 - \frac{1}{3}z^3 \frac{2i-n-1}{\sqrt{(i-1)(n-i)(n-1)}}\right\} + \mathcal{O}(n^{-1+\frac{4}{5}\varepsilon})$$

uniformly for $|z| \leqq n^{\frac{1}{5}\varepsilon}$, $0 < \varepsilon < \frac{5}{4}$. We note that the fact that $h(z) = 0$ outside the interval $-\sqrt{\frac{(i-1)(n-1)}{n-i}} < z < \sqrt{\frac{(n-i)(n-1)}{i-1}}$ is of no consequence as the set $|z| \leqq n^{\frac{1}{5}\varepsilon}$ is contained in this interval for sufficiently large values of n.

From the MARKOV inequality (cf. [21], 158) we find

$$P(|\underline{z}| > n^{\frac{1}{5}\varepsilon}) \leqq n^{-\frac{2}{5}j\varepsilon} E\, \underline{z}^{2j} = \mathcal{O}(n^{-1}) ,$$

by choosing $j \geqq \frac{5}{2\varepsilon}$, since the moments of $\underline{z}$ are bounded as n tends to infinity. By a similar argument one easily shows

$$\int_{|t| > n^{\frac{1}{5}\varepsilon}} |\bar{h}(t)|\, dt = \mathcal{O}(n^{-1}) ,$$

where

$$\bar{h}(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} \left\{ 1 - \frac{1}{3}t^3 \frac{2i-n-1}{\sqrt{(i-1)(n-i)(n-1)}} \right\}$$

Hence for the distribution function H of $\underline{z}$ we have

$$H(z) = \int_{-\infty}^{z} h(t)\,dt = \int_{-n^{\frac{1}{5}\varepsilon}}^{\min(z,n^{\frac{1}{5}\varepsilon})} h(t)\,dt + \mathcal{O}(n^{-1}) =$$

$$= \int_{-n^{\frac{1}{5}\varepsilon}}^{\min(z,n^{\frac{1}{5}\varepsilon})} \bar{h}(t)\,dt + \mathcal{O}(n^{-1+\frac{4}{5}\varepsilon}) \int_{-n^{\frac{1}{5}\varepsilon}}^{\min(z,n^{\frac{1}{5}\varepsilon})} dt + \mathcal{O}(n^{-1}) =$$

$$= \int_{-\infty}^{z} \bar{h}(t)\,dt + \mathcal{O}(n^{-(1-\varepsilon)}) =$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{1}{2}t^2}\,dt + \frac{1}{3} \frac{2i-n-1}{\sqrt{(i-1)(n-i)(n-1)}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} (z^2+2) + \mathcal{O}(n^{-(1-\varepsilon)})$$

for all values of z and any $0 < \varepsilon < \frac{5}{4}$, hence any $\varepsilon > 0$.

To find an asymptotic expression for the median $m(\underline{z})$ of $\underline{z}$ we have to solve the equation

$$H(m(\underline{z})) = \frac{1}{2} .$$

Since it is clear that $m(\underline{z})$ tends to zero as n tends to infinity we write

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{m(\underline{z})} e^{-\frac{1}{2}t^2}\,dt = \frac{1}{2} + \frac{1}{\sqrt{2\pi}} m(\underline{z})\, e^{-\frac{1}{2}m^2(\underline{z})} + \mathcal{O}(m^3(\underline{z})) ,$$

and solve

$$\frac{1}{2} + \frac{1}{\sqrt{2\pi}} m(\underline{z})\, e^{-\frac{1}{2}m^2(\underline{z})} + \frac{1}{3}\frac{2i-n-1}{\sqrt{(i-1)(n-i)(n-1)}}\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}m^2(\underline{z})}(m^2(\underline{z})+2) +$$

$$+\ \mathcal{O}(m^3(\underline{z})) + \mathcal{O}(n^{-(1-\varepsilon)}) = \frac{1}{2}\ ,$$

or

$$m(\underline{z}) + \mathcal{O}(m^2(\underline{z})) = -\frac{2}{3}\frac{2i-n-1}{\sqrt{(i-1)(n-i)(n-1)}} + \mathcal{O}(n^{-(1-\varepsilon)})\ ,$$

or

$$m(\underline{z}) = -\frac{2}{3}\frac{2i-n-1}{\sqrt{(i-1)(n-i)(n-1)}} + \mathcal{O}(n^{-(1-\varepsilon)})\ .$$

We note that this expression for $m(\underline{z})$ is valid even if $(2i-n-1)n^{-\frac{1}{2}}$ is bounded in which case the first term need not be the leading term. For the median $m(i,n)$ of $\underline{y}$ we now have

$$m(i,n) = m(\underline{z})\sqrt{\frac{(i-1)(n-i)}{(n-1)^3}} + \frac{i-1}{n-1} =$$

$$= \frac{i-\frac{1}{3}}{n+\frac{1}{3}} + \mathcal{O}(n^{-\frac{3}{2}+\varepsilon})$$

for all $\varepsilon > 0$. This completes the proof of lemma 3.2.1.

It may be of interest to note that the expression obtained for $m(i,n)$ is equivalent to the statement that for $n \to \infty$, $\frac{i}{n} \to r$, $0 < r < 1$, $r \neq \frac{1}{2}$,

$$\frac{m(i,n) - M(i,n)}{\mu(i,n) - M(i,n)} = \frac{2}{3} + \mathcal{O}(n^{-\frac{1}{2}+\varepsilon})\ ,$$

where $\mu(i,n) = \frac{i}{n+1}$ and $M(i,n) = \frac{i-1}{n-1}$ denote expectation and mode of the distribution $B_{i:n}$. This is the famous rule of thumb

$$(\text{median} - \text{mode}) \approx \frac{2}{3}(\text{mean} - \text{mode})$$

that has been discussed by K. PEARSON [22], J.B.S. HALDANE [12] and M.G. KENDALL and A. STUART [20]. One might therefore try and prove the result by using the theorem quoted by KENDALL and STUART ([20], 179) in this connection; one would then have to show that for $k \geq 2$ the k-th cumulant of the distribution $B_{i:n}$ is $\mathcal{O}(n^{1-k})$ as i and n tend to infinity under the conditions given.

Making use of this lemma we proceed to obtain asymptotic expressions for $E\,\underline{x}_{i:n}$, $\mu_k(\underline{x}_{i:n})$ and $m(\underline{x}_{i:n})$. The results are given in lemmata 3.2.2, 3.2.3 and 3.2.4.

<u>LEMMA 3.2.2</u>

If $F \in \mathcal{F}$ and $\lim \frac{i}{n} = r$, $0 < r < 1$, as n tends to infinity, then for sufficiently large n $E\,\underline{x}_{i:n}$ exists and

$$E\,\underline{x}_{i:n} = G(\tfrac{i}{n+1}) + \tfrac{1}{2} G''(\tfrac{i}{n+1})\,\mu_2(i,n) + o(n^{-1}) \quad .$$

<u>PROOF</u>

Let $\varepsilon_n = \max \left\{ \left|\frac{i}{n+1} - r\right|, 2\left|\frac{i-a}{n-a-b+1} - r\right|, \frac{1}{\log (n+1)} \right\}$ for $n \geq a+b$,

where a and b denote the constants of (3.1.11). Clearly $\varepsilon_n > 0$, $\lim_{n\to\infty} \varepsilon_n = 0$ and $\varepsilon_n < \frac{r}{2}$ and $\varepsilon_n < \frac{1-r}{2}$ for sufficiently large n, say $n \geq N$. By (3.1.11)

$$|G(y) b_{i:n}(y)| = \left|\frac{n!(i-a-1)!(n-i-b)!}{(n-a-b)!(i-1)!(n-i)!} G(y)\, y^a\, (1-y)^b\, b_{i-a:n-a-b}(y)\right| \leq$$

$$\leq M.b_{i-a:n-a-b}(y) \qquad \text{for } 0 < y < 1 \quad ,$$

where M is a constant independent of i, n and y. We note that for $n \geq N$, $i-a > 0$ and $n-i-b > 0$, and hence $E\,\underline{x}_{i:n}$ exists.

Hence we have for $n \geq N$, and M' independent of i and n,

$$\left| \int_{|y-r| > \varepsilon_n} \{G(y) - G(\tfrac{i}{n+1}) - G'(\tfrac{i}{n+1})(y - \tfrac{i}{n+1}) - \tfrac{1}{2}G''(\tfrac{i}{n+1})(y-\tfrac{i}{n+1})^2\} \cdot b_{i:n}(y)\, dy \right| \leq$$

$$\leq M' \int_{|y-r| > \varepsilon_n} b_{i-a:n-a-b}(y)\, dy \leq$$

$$\leq M' \int_{|y - \frac{i-a}{n-a-b+1}| > \frac{1}{2}\varepsilon_n} b_{i-a:n-a-b}(y)\, dy \leq$$

$$\leq M' \frac{\mu_4(i-a,n-a-b)}{(\frac{1}{2}\varepsilon_n)^4} \leq 16\, M'\, \mu_4(i-a,n-a-b)(\log\,(n+1))^4 = \mathcal{O}(n^{-1})$$

by MARKOV's inequality and lemma 3.2.1.

Writing

$$(3.2.2) \quad G(y) = G(\tfrac{i}{n+1}) + G'(\tfrac{i}{n+1})(y - \tfrac{i}{n+1}) + \tfrac{1}{2}G''(\theta y + (1-\theta)\tfrac{i}{n+1})(y - \tfrac{i}{n+1})^2$$

where $0 \leq \theta \leq 1$ and θ may depend on y, we have for $n \geq N$,

$$\left| \int_{|y-r| \leq \varepsilon_n} \{G(y) - G(\tfrac{i}{n+1}) - G'(\tfrac{i}{n+1})(y - \tfrac{i}{n+1}) - \tfrac{1}{2}G''(\tfrac{i}{n+1})(y-\tfrac{i}{n+1})^2\} \cdot b_{i:n}(y)\, dy \right| =$$

$$= \left| \frac{1}{2} \int_{|y-r| \leq \varepsilon_n} \{G''(\theta y + (1-\theta)\tfrac{i}{n+1}) - G''(\tfrac{i}{n+1})\} (y - \tfrac{i}{n+1})^2\, b_{i:n}(y)\, dy \right| \leq$$

$$\leq \frac{1}{2} \sup_{\substack{r-\varepsilon_n < y_1 < r+\varepsilon_n \\ r-\varepsilon_n < y_2 < r+\varepsilon_n}} |G''(y_1) - G''(y_2)|\, \mu_2(i,n) = \mathcal{O}(n^{-1}) \quad ,$$

because of the continuity of G'' and lemma 3.2.1. Hence

$$E\,\underline{x}_{i:n} - G(\tfrac{i}{n+1}) - \tfrac{1}{2}G''(\tfrac{i}{n+1})\,\mu_2(i,n) =$$

$$= \int_0^1 \{G(y) - G(\tfrac{i}{n+1}) - G'(\tfrac{i}{n+1})(y-\tfrac{i}{n+1}) - \tfrac{1}{2}G''(\tfrac{i}{n+1})(y-\tfrac{i}{n+1})^2\} \,.$$

$$. \; b_{i:n}(y)\,dy =$$

$$= \mathcal{O}(n^{-1}) \;,$$

which proves the lemma.

We remark that in general if G is 2p-times continuously differentiable one proves in the same way that

$$E\,\underline{x}_{i:n} = \sum_{j=0}^{2p} \frac{1}{j!}\, G^{(j)}(\tfrac{i}{n+1})\, \mu_j(i,n) + \mathcal{O}(n^{-p}) \quad .$$

<u>LEMMA 3.2.3</u>

If $F \in \mathcal{F}$, $k = 2,3,\ldots$, and $\lim \frac{i}{n} = r$, $0 < r < 1$, as n tends to infinity, then for sufficiently large n $\mu_k(\underline{x}_{i:n})$ exists and

$$\mu_k(\underline{x}_{i:n}) = (G'(\tfrac{i}{n+1}))^k\, \mu_k(i,n) + \mathcal{O}(n^{-\frac{k+1}{2}}) \qquad \text{if } k \text{ is even ,}$$

$$\mu_k(\underline{x}_{i:n}) = (G'(\tfrac{i}{n+1}))^k\, \mu_k(i,n) + \frac{k}{2}\,(G'(\tfrac{i}{n+1}))^{k-1}\, G''(\tfrac{i}{n+1}) \,.$$

$$. \left[\mu_{k+1}(i,n) - \mu_2(i,n)\mu_{k-1}(i,n)\right] + \mathcal{O}(n^{-\frac{k+1}{2}}) \text{ if } k \text{ is odd .}$$

<u>PROOF</u>

Let $\varepsilon_n = \max\{|\frac{i}{n+1} - r|,\ 2|\frac{i-ka}{n-ka-kb+1} - r|,\ \frac{1}{\log\,(n+1)}\}$, $n \geq ka+kb$,

for some fixed $k = 2,3,\ldots$, where a and b denote the constants in (3.1.11). Clearly $\varepsilon_n > 0$, $\lim_{n\to\infty} \varepsilon_n = 0$ and for $n \geq N$, $\varepsilon_n < \frac{r}{2}$ and $\varepsilon_n < \frac{1-r}{2}$. By (3.1.11)

$$|G^k(y)\, b_{i:n}(y)| = \left|\frac{n!(i-ka-1)!(n-i-kb)!}{(n-ka-kb)!(i-1)!(n-i)!}\, G^k(y)\, y^{ka}\, (1-y)^{kb}\, .\right.$$

$$\left.\cdot\, b_{i-ka:n-ka-kb}(y)\right| \leqq$$

$$\leqq M\, b_{i-ka:n-ka-kb}(y) \qquad \text{for } 0 < y < 1 \quad ,$$

where M is independent of i, n and y. For $n \geqq N$, $i-ka > 0$ and $n-i-kb > 0$, and hence $\mu_k(\underline{x}_{i:n})$ exists.

Thus for $n \geqq N$, and M' independent of i and n,

(3.2.3)
$$\left| \int_{|y-r| > \varepsilon_n} (G(y) - E\,\underline{x}_{i:n})^k\, b_{i:n}(y)\, dy \right| \leqq$$

$$\leqq M' \int_{|y-r| > \varepsilon_n} b_{i-ka:n-ka-kb}(y)\, dy \leqq$$

$$\leqq M' \int_{|y - \frac{i-ka}{n-ka-kb+1}| > \frac{1}{2}\varepsilon_n} b_{i-ka:n-ka-kb}(y)\, dy \leqq$$

$$\leqq M' \frac{\mu_{2k+2}(i-ka, n-ka-kb)}{(\frac{1}{2}\varepsilon_n)^{2k+2}} \leqq$$

$$\leqq 2^{2k+2}\, M'\, \mu_{2k+2}(i-ka, n-ka-kb)(\log\,(n+1))^{2k+2} =$$

$$= \mathcal{O}(n^{-k-\frac{1}{2}}) = \mathcal{O}(n^{-\frac{k+1}{2}}) \quad ,$$

by MARKOV's inequality and lemma 3.2.1. Also by MARKOV's inequality, one may show that for any $l \geqq 0$

(3.2.4)
$$\left| \int_{|y-r| > \varepsilon_n} (y - \tfrac{i}{n+1})^l\, b_{i:n}(y)\, dy \right| = \mathcal{O}(n^{-\frac{k+1}{2}}) \quad .$$

Using lemma 3.2.2 and (3.2.2) we have

$$(3.2.5)\quad \int_{|y-r|\leq\varepsilon_n} (G(y) - E\,\underline{x}_{i:n})^k\, b_{i:n}(y)\,dy =$$

$$= \int_{|y-r|\leq\varepsilon_n} \Big\{G'(\tfrac{i}{n+1})(y-\tfrac{i}{n+1}) + \tfrac{1}{2}G''(\theta y + (1-\theta)\tfrac{i}{n+1})(y-\tfrac{i}{n+1})^2 + $$

$$- \tfrac{1}{2}G''(\tfrac{i}{n+1})\,\mu_2(i,n) + \mathcal{O}(n^{-1})\Big\}^k\, b_{i:n}(y)\,dy =$$

$$= \sum{}^{*} \frac{k!}{k_1!k_2!\ldots k_5!}\,(G'(\tfrac{i}{n+1}))^{k_1}(\tfrac{1}{2}G''(\tfrac{i}{n+1}))^{k_2+k_3}(-\mu_2(i,n))^{k_3}\,.$$

$$.\;(\mathcal{O}(n^{-1}))^{k_5}\int_{|y-r|\leq\varepsilon_n}\Big\{\tfrac{1}{2}G''(\theta y + (1-\theta)\tfrac{i}{n+1}) - \tfrac{1}{2}G''(\tfrac{i}{n+1})\Big\}^{k_4}\,.$$

$$.\;(y-\tfrac{i}{n+1})^{k_1+2k_2+2k_4}\, b_{i:n}(y)\,dy\;,$$

where the latter expression is obtained by writing

$$G''(\theta y + (1-\theta)\tfrac{i}{n+1}) = G''(\tfrac{i}{n+1}) + \Big\{G''(\theta y + (1-\theta)\tfrac{i}{n+1}) - G''(\tfrac{i}{n+1})\Big\}$$

and expanding the multinomial in the integrand; by $\sum^{*}$ we indicate that the summation ranges over all non-negative integers $k_1, k_2, \ldots, k_5$ satisfying $\sum_{j=1}^{5} k_j = k$.

Consider a term T in (3.2.5) corresponding to a set of values for $k_1, k_2, \ldots, k_5$ having $k_4 \neq 0$. By lemma 3.2.1 for some constant C

$$|T| \leq C.n^{-k_3-k_5} \sup_{\substack{r-\varepsilon_n < y_1 < r+\varepsilon_n \\ r-\varepsilon_n < y_2 < r+\varepsilon_n}} |G''(y_1) - G''(y_2)|^{k_4}\,.$$

$$.\int_0^1 |y-\tfrac{i}{n+1}|^{k_1+2k_2+2k_4}\, b_{i:n}(y)\,dy =$$

$$= \mathcal{O}(n^{-\frac{1}{2}k_1-k_2-k_3-k_4-k_5}) = \mathcal{O}(n^{-\frac{k+1}{2}}) ,$$

since G" is continuous and $\lim_{n\to\infty} \varepsilon_n = 0$.

Also, if $k_5 \neq 0$ we have

$$T = \mathcal{O}(n^{-k_5}) . \mathcal{O}(n^{-k_3-\frac{k_1}{2}-k_2-k_4}) = \mathcal{O}(n^{-\frac{k+1}{2}}) ,$$

and hence from (3.2.3) and (3.2.5)

$$(3.2.6) \quad \mu_k(\underline{x}_{i:n}) = \int_0^1 (G(y) - E\,\underline{x}_{i:n})^k\, b_{i:n}(y)\, dy =$$

$$= \sum{}^{*} \frac{k!}{k_1!k_2!k_3!} \left(G'\left(\frac{i}{n+1}\right)\right)^{k_1} \left(\frac{1}{2}G''\left(\frac{i}{n+1}\right)\right)^{k_2+k_3} (-\mu_2(i,n))^{k_3} .$$

$$. \int_{|y-r| \leq \varepsilon_n} \left(y - \frac{i}{n+1}\right)^{k_1+2k_2} b_{i:n}(y)\, dy + \mathcal{O}(n^{-\frac{k+1}{2}}) =$$

$$= \sum{}^{*} \frac{k!}{k_1!k_2!k_3!} \left(G'\left(\frac{i}{n+1}\right)\right)^{k_1} \left(\frac{1}{2}G''\left(\frac{i}{n+1}\right)\right)^{k_2+k_3} (-\mu_2(i,n))^{k_3} .$$

$$. \mu_{k_1+2k_2}(i,n) + \mathcal{O}(n^{-\frac{k+1}{2}})$$

by (3.2.4). Here $\sum^{*}$ denotes summation over all non-negative integers k_1, k_2, k_3 satisfying $k_1 + k_2 + k_3 = k$.

To every triplet k_1, k_2, k_3, with $\sum_{j=1}^{3} k_j = k$, there corresponds a term T in (3.2.6) and by lemma 3.2.1

$$T = \mathcal{O}\left(n^{-\left[\frac{k_1+1}{2}\right] -k_2-k_3}\right) .$$

For even values of k and $k \neq k_1$

$$T = \mathcal{O}(n^{-\frac{k}{2}-1}) = \mathcal{O}(n^{-\frac{k+1}{2}}) , \quad \text{hence}$$

$$\mu_k(\underline{x}_{i:n}) = (G'(\tfrac{i}{n+1}))^k \mu_k(i,n) + \mathcal{O}(n^{-\frac{k+1}{2}}) \qquad \text{for even } k .$$

For odd values of k and $k_1 \leqq k-2$ we have

$$T = \mathcal{O}(n^{-\frac{k+3}{2}}) = \mathcal{O}(n^{-\frac{k+1}{2}}) , \quad \text{hence}$$

$$\mu_k(\underline{x}_{i:n}) = (G'(\tfrac{i}{n+1}))^k \mu_k(i,n) + \frac{k}{2} (G'(\tfrac{i}{n+1}))^{k-1} G''(\tfrac{i}{n+1}) \,.$$

$$. \left[\mu_{k+1}(i,n) - \mu_2(i,n)\mu_{k-1}(i,n)\right] + \mathcal{O}(n^{-\frac{k+1}{2}}) \quad \text{for odd } k .$$

This completes the proof of lemma 3.2.3. We note that from this result one easily proves the asymptotic normality of $\underline{x}_{i:n}$ by moment convergence.

<u>LEMMA 3.2.4</u>

If $F \in \mathcal{F}$ and $\lim \frac{i}{n} = r$, $0 < r < 1$, as n tends to infinity, then for all $\varepsilon > 0$

$$m(\underline{x}_{i:n}) = G\left(\frac{i - \frac{1}{3}}{n + \frac{1}{3}}\right) + \mathcal{O}(n^{-\frac{3}{2}+\varepsilon}) =$$

$$= G(\tfrac{i}{n+1}) + \frac{1}{3} \frac{2i-n-1}{(n+1)^2} G'(\tfrac{i}{n+1}) + \mathcal{O}(n^{-\frac{3}{2}+\varepsilon}) .$$

<u>PROOF</u>

From (3.1.1) and lemma 3.2.1 we have

$$(3.2.7) \qquad F(m(\underline{x}_{i:n})) = m(i,n) = \frac{i-\frac{1}{3}}{n+\frac{1}{3}} + \mathcal{O}(n^{-\frac{3}{2}+\varepsilon}) \;,$$

or

$$m(\underline{x}_{i:n}) = G(m(i,n)) = G\left(\frac{i-\frac{1}{3}}{n+\frac{1}{3}} + \mathcal{O}(n^{-\frac{3}{2}+\varepsilon})\right) =$$

$$= G\left(\frac{i-\frac{1}{3}}{n+\frac{1}{3}}\right) + G'\left(\frac{i-\frac{1}{3}}{n+\frac{1}{3}} + \theta\,\mathcal{O}(n^{-\frac{3}{2}+\varepsilon})\right).\,\mathcal{O}(n^{-\frac{3}{2}+\varepsilon}) =$$

$$= G\left(\frac{i-\frac{1}{3}}{n+\frac{1}{3}}\right) + \mathcal{O}(n^{-\frac{3}{2}+\varepsilon}) \;,$$

since G' is continuous on (0,1),

$$\lim \left(\frac{i-\frac{1}{3}}{n+\frac{1}{3}} + \theta\,\mathcal{O}(n^{-\frac{3}{2}+\varepsilon})\right) = r \;, \qquad \text{for } 0 < \varepsilon < \frac{3}{2} \;,$$

r is an interior point of (0,1), and hence $G'\left(\frac{i-\frac{1}{3}}{n+\frac{1}{3}} + \theta\,\mathcal{O}(n^{-\frac{3}{2}+\varepsilon})\right)$ is bounded. Alternatively we may write

$$G\left(\frac{i-\frac{1}{3}}{n+\frac{1}{3}}\right) = G\left(\frac{i}{n+1} + \frac{1}{3}\,\frac{2i-n-1}{(n+1)^2} + \mathcal{O}(n^{-2})\right) =$$

$$= G(\frac{i}{n+1}) + \{\frac{1}{3}\,\frac{2i-n-1}{(n+1)^2} + \mathcal{O}(n^{-2})\}\,G'(\frac{i}{n+1}) + \mathcal{O}(n^{-2}) =$$

$$= G(\frac{i}{n+1}) + \frac{1}{3}\,\frac{2i-n-1}{(n+1)^2}\,G'(\frac{i}{n+1}) + \mathcal{O}(n^{-2}) \;,$$

since G' and G'' are continuous on (0,1) and $0 < r < 1$. This proves the lemma.

We are now in a position to prove three theorems that will be needed in the next chapter.

THEOREM 3.2.1

If $F \in \mathcal{F}$ and $\lim \frac{i}{n} = r$, $0 < r < 1$, as n tends to infinity, then $F(E\,\underline{x}_{i:n})$ exists for sufficiently large n, and

$$F(E\,\underline{x}_{i:n}) = \frac{i}{n+1} + \frac{i(n+1-i)}{2(n+1)^3} \frac{G''(\frac{i}{n+1})}{G'(\frac{i}{n+1})} + \mathcal{O}(n^{-1}) \quad .$$

PROOF

From lemma 3.2.2 we have

$$F(E\,\underline{x}_{i:n}) = F\Big(G(\tfrac{i}{n+1}) + \tfrac{1}{2}G''(\tfrac{i}{n+1})\ \mu_2(i,n) + \mathcal{O}(n^{-1})\Big) =$$

$$= FG(\tfrac{i}{n+1}) + F'G(\tfrac{i}{n+1})\Big(\tfrac{1}{2}G''(\tfrac{i}{n+1})\ \mu_2(i,n) + \mathcal{O}(n^{-1})\Big) +$$

$$+ F''\Big(\theta G(\tfrac{i}{n+1}) + (1-\theta)E\,\underline{x}_{i:n}\Big)\mathcal{O}(n^{-2}) \quad ,$$

since F'' is continuous on I by (3.1.6). As by lemma 3.2.2

$$\lim_{n\to\infty} E\,\underline{x}_{i:n} = \lim_{n\to\infty} G(\tfrac{i}{n+1}) = G(r)$$

and G(r) is an interior point of I, $F'G(\frac{i}{n+1})$ and $F''\Big(\theta G(\frac{i}{n+1}) + (1-\theta)E\,\underline{x}_{i:n}\Big)$ are bounded. Hence

$$F(E\,\underline{x}_{i:n}) = FG(\tfrac{i}{n+1}) + \tfrac{1}{2}F'G(\tfrac{i}{n+1})\ G''(\tfrac{i}{n+1})\ \mu_2(i,n) + \mathcal{O}(n^{-1}) =$$

$$= \frac{i}{n+1} + \frac{i(n+1-i)}{2(n+1)^3} \frac{G''(\frac{i}{n+1})}{G'(\frac{i}{n+1})} + \mathcal{O}(n^{-1}) \quad ,$$

by (3.1.3) and (3.2.1) and since $F'G(y).G'(y) \equiv (FG(y))' \equiv 1$, or $F'G(y) \equiv \frac{1}{G'(y)}$ for $0 < y < 1$.

THEOREM 3.2.2

If $F \in \mathcal{T}$, $k = 1,2,\ldots$, and $\lim \frac{i}{n} = r$, $0 < r < 1$, as n tends to infinity, then for sufficiently large n $\frac{\mu_{2k+1}(\underline{x}_{i:n})}{\sigma^{2k+1}(\underline{x}_{i:n})}$ exists and

$$\frac{\mu_{2k+1}(\underline{x}_{i:n})}{\sigma^{2k+1}(\underline{x}_{i:n})} = \frac{\mu_{2k+1}(i,n)}{(\mu_2(i,n))^{k+\frac{1}{2}}} + 2^{-k}\,\frac{(2k+1)!}{(k-1)!}\left[\frac{i(n+1-i)}{(n+1)^3}\right]^{\frac{1}{2}}\frac{G''(\frac{i}{n+1})}{G'(\frac{i}{n+1})} + \mathcal{O}(n^{-\frac{1}{2}}).$$

PROOF

The theorem follows from lemmata 3.2.1 and 3.2.3 by straightforward algebra.

THEOREM 3.2.3

If $F \in \mathcal{T}$ and $\lim \frac{i}{n} = r$, $0 < r < 1$, as n tends to infinity, then for sufficiently large n $\frac{E\,\underline{x}_{i:n} - m(\underline{x}_{i:n})}{\sigma(\underline{x}_{i:n})}$ exists and

$$\frac{E\,\underline{x}_{i:n} - m(\underline{x}_{i:n})}{\sigma(\underline{x}_{i:n})} = -\frac{2i-n-1}{3\sqrt{i(n+1-i)(n+1)}} + \frac{1}{2}\left[\frac{i(n+1-i)}{(n+1)^3}\right]^{\frac{1}{2}}\frac{G''(\frac{i}{n+1})}{G'(\frac{i}{n+1})} + \mathcal{O}(n^{-\frac{1}{2}}).$$

PROOF

The proof is immediate upon applying lemmata 3.2.2 and 3.2.4.

It may be of interest to compare the results of lemma 3.2.2 and theorem 3.2.1 with the corresponding results obtained by G. BLOM [4], that we mentioned at the beginning of this section. To this end we write

$$\text{(3.2.8)} \qquad E\,\underline{x}_{i:n} = G\left(\frac{i-\alpha_{i:n}}{n+1-\alpha_{i:n}-\beta_{i:n}}\right) , \qquad \text{or}$$

$$\text{(3.2.8')} \qquad F(E\,\underline{x}_{i:n}) = \frac{i-\alpha_{i:n}}{n+1-\alpha_{i:n}-\beta_{i:n}} .$$

Although these equations do not determine $\alpha_{i:n}$ and $\beta_{i:n}$ uniquely we have from theorem 3.2.1 the following

COROLLARY 3.2.1

If $F \in \mathcal{F}$ and $\lim \frac{i}{n} = r$, $0 < r < 1$, as n tends to infinity, and if, in addition, $\alpha(r) = \lim \alpha_{i:n}$ and $\beta(r) = \lim \beta_{i:n}$ exist, then $\alpha(r)$ and $\beta(r)$ satisfy

$$\frac{1}{2}r(1-r)G''(r) + \left[\alpha(r)(1-r) - \beta(r)r\right] G'(r) = 0 .$$

PROOF

Since $\alpha_{i:n}$ and $\beta_{i:n}$ tend to finite limits they remain bounded as n tends to infinity. From theorem 3.2.1 we know that for sufficiently large n $F(E\,\underline{x}_{i:n})$ exists, and hence

$$F(E\,\underline{x}_{i:n}) = \frac{i-\alpha_{i:n}}{n+1-\alpha_{i:n}-\beta_{i:n}} = \frac{i}{n+1}\left(1-\frac{\alpha_{i:n}}{i}\right)\left(1-\frac{\alpha_{i:n}+\beta_{i:n}}{n+1}\right)^{-1} =$$

$$= \frac{i}{n+1} - \frac{\alpha_{i:n}}{n+1} + \frac{i}{(n+1)^2}(\alpha_{i:n}+\beta_{i:n}) + \mathcal{O}(n^{-2}) =$$

$$= \frac{i}{n+1} - \alpha_{i:n}\frac{n+1-i}{(n+1)^2} + \beta_{i:n}\frac{i}{(n+1)^2} + \mathcal{O}(n^{-2}) =$$

$$= \frac{i}{n+1} + \frac{i(n+1-i)}{2(n+1)^3}\,\frac{G''(\frac{i}{n+1})}{G'(\frac{i}{n+1})} + \mathcal{O}(n^{-1}) ,$$

by theorem 3.2.1. Multiplying the last two members by (n+1) and letting n tend to infinity leads to the desired result.

Corollary 3.2.1 is precisely BLOM's result ([4], 65). The only major difference is that BLOM imposes more restrictive regularity conditions on the higher derivatives of G in order to obtain more information about the order of the remainder term in his version of lemma 3.2.2. This also enables him to consider the case that $\frac{i}{n}$ tends to 0 or 1, provided the convergence is sufficiently slow.

In order to determine $\alpha_{i:n}$ and $\beta_{i:n}$ uniquely when $F(E\,\underline{x}_{i:n})$ exists and $i \neq \frac{n+1}{2}$ BLOM adds the restrictions

(3.2.9) $$\alpha_{i:n} = \alpha_{n+1-i:n} \quad , \qquad \beta_{i:n} = \beta_{n+1-i:n} \quad .$$

Under the conditions of corollary 3.2.1 we then find for $r \neq \frac{1}{2}$

(3.2.10) $$\alpha(r) = \alpha(1-r) = \frac{r(1-r)}{2(2r-1)} \left[(1-r)\,\frac{G''(r)}{G'(r)} - r\,\frac{G''(1-r)}{G'(1-r)}\right]$$
$$\beta(r) = \beta(1-r) = \frac{r(1-r)}{2(2r-1)} \left[r\,\frac{G''(r)}{G'(r)} - (1-r)\,\frac{G''(1-r)}{G'(1-r)}\right] \quad ,$$

whereas for $r = \frac{1}{2}$, $\alpha(r)$ and $\beta(r)$ are indeterminate.

In the special case that F denotes a symmetric distribution we have $F(E\,\underline{x}_{i:n}) = 1 - F(E\,\underline{x}_{n+1-i:n})$ or $\alpha_{i:n} = \beta_{i:n}$, and $G'(y) = G'(1-y)$ and $G''(y) = -G''(1-y)$. Hence in the symmetrical case

(3.2.11) $$F(E\,\underline{x}_{i:n}) = \frac{i-\alpha_{i:n}}{n+1-2\alpha_{i:n}} \quad ,$$

and under the conditions of corollary 3.2.1, for $r \neq \frac{1}{2}$,

(3.2.12) $$\alpha(r) = \alpha(1-r) = \frac{r(1-r)}{2(2r-1)}\,\frac{G''(r)}{G'(r)} \quad .$$

An important part of BLOM's research is devoted to the study of the behaviour of $\alpha(r)$ and $\beta(r)$ for $\frac{1}{2} < r < 1$ for specific distributions F. It turns out that in general (cf. [4], 76, or chapter 5 of the present study for the exception to this rule) $\alpha(r)$ and $\beta(r)$ do indeed depend on r and hence that even if n tends to infinity, there are no uniformly best values for α and β for all i. It is, however, possible to determine the range of $\alpha(r)$ and $\beta(r)$ for $\frac{1}{2} < r < 1$ and construct asymptotic upper and lower bounds for $F(E\ \underline{x}_{i:n})$ for specific distributions F in this way. In chapter 5 we shall incidentally generalize one of these inequalities to finite sample sizes.

In contrast to the difficulties arising in studying the asymptotic properties of $\alpha_{i:n}$ and $\beta_{i:n}$ in (3.2.8) the analogous study for the median turns out to be exceedingly simple. Writing

$$(3.2.13) \qquad m(\underline{x}_{i:n}) = G\left(\frac{i-\alpha'_{i:n}}{n+1-\alpha'_{i:n}-\beta'_{i:n}}\right) \quad , \quad \text{or}$$

$$(3.2.13') \qquad F(m(\underline{x}_{i:n})) = \frac{i-\alpha'_{i:n}}{n+1-\alpha'_{i:n}-\beta'_{i:n}} \quad , \quad \text{with}$$

$$(3.2.14) \qquad \alpha'_{i:n} = \alpha'_{n+1-i:n} \quad , \quad \beta'_{i:n} = \beta'_{n+1-i:n} \quad ,$$

we have from (3.1.1) $F(m(\underline{x}_{n+1-i:n})) = m(n+1-i,n) = 1 - m(i,n) =$ $= 1 - F(m(\underline{x}_{i:n}))$ and hence $\alpha'_{i:n} = \beta'_{i:n}$ for all i and n and any distribution F, or

$$(3.2.15) \qquad F(m(\underline{x}_{i:n})) = \frac{i-\alpha'_{i:n}}{n+1-2\alpha'_{i:n}} \quad ,$$

where $\alpha'_{i:n}$ does not depend on F.
Furthermore we have from lemma 3.2.4 that if $F \in \mathcal{F}$ and $\frac{i}{n}$ tends to r,

$0 < r < 1$, $r \neq \frac{1}{2}$, then $\lim \alpha'_{i:n} = \alpha'(r)$, with

$$(3.2.16) \qquad \alpha'(r) = \alpha'(1-r) = \frac{1}{3} .$$

For $r = \frac{1}{2}$ one has the usual indeterminacy.

This property of $m(\underline{x}_{i:n})$ has some significance for the approach to the problem of plotting on probability papers devised by A. BENARD and E.C. BOS-LEVENBACH [2]. This problem may be explained as follows (cf. also [4], 143).

Suppose we have a random sample $\underline{z}_1, \underline{z}_2, \dots, \underline{z}_n$ which is thought to have been drawn from a continuous distribution $F(\frac{z-\mu}{\sigma})$, where F is a known distribution and μ and σ are unknown location and scale parameters. To verify the assumption that the distribution is indeed of type F and to estimate μ and σ graphically, one often makes use of the appropriate probability paper which has its vertical scale proportional to G, so that the points $(z, F(\frac{z-\mu}{\sigma}))$ form a straight line of which the height and slope are determined by μ and σ. After ordering the sample $z_{i:n} < z_{2:n} < \dots < z_{n:n}$ one plots the points $(z_{i:n}, P_i)$ on this probability paper, where P_i is some estimate of $F\left(\frac{z_{i:n} - \mu}{\sigma}\right)$. Then one draws a straight line through these n points by hand and estimates μ and σ from the height and slope of this line; non-linearity of the points is evidence that the parent distribution is not of type F, as was assumed.

The problem is of course the determination of the quantities P_i. Commonly advocated choices for all F seem to be $\frac{i}{n}$, $\frac{i-1}{n}$, $\frac{i-\frac{1}{2}}{n}$, $\frac{i}{n+1}$ and $\frac{i-1}{n-1}$ (cf. [2]); a sensible approach to the problem seems to be the following. The choice

$$(3.2.17) \qquad P_i = F\left(\frac{E\,\underline{z}_{i:n} - \mu}{\sigma}\right) = F(E\,\underline{x}_{i:n}) = \frac{i-\alpha_{i:n}}{n+1-\alpha_{i:n}-\beta_{i:n}} ,$$

where $\underline{x}_{i:n}$ denotes the i-th order statistic of a sample of size n from the distribution F and $\alpha_{i:n}$ and $\beta_{i:n}$ are defined by (3.2.8) and

(3.2.9) for F, leads to plotting points $\left(\underline{z}_{i:n}, F\left(\frac{E\,\underline{z}_{i:n} - \mu}{\sigma}\right)\right)$ (where only the first coordinate is random) to estimate the point $\left(E\,\underline{z}_{i:n}, F\left(\frac{E\,\underline{z}_{i:n} - \mu}{\sigma}\right)\right)$ on the line $(z, F(\frac{z-\mu}{\sigma}))$, and hence one estimates this line at n points by an unbiased estimator. Since for a "quick and dirty" method like this different values of $\alpha_{i:n}$ and $\beta_{i:n}$ for every i and n are not readily acceptable, this line of thought has led BLOM ([4], 145) to propose for the normal distribution an "intermediate" value $\alpha = \beta = \frac{3}{8}$, or

$$(3.2.18) \qquad P_i = \frac{i-\frac{3}{8}}{n+1-\frac{3}{4}} .$$

This proposal may, however, be objected to for the following reasons:

1. The property of unbiasedness is only appealing if one assumes something like "the fitted line is identical with the line obtained by minimizing the sum of squares of the horizontal deviations from the line" (cf. [4], 143, [7]).
2. Even the "intermediate" values for α and β (which for the normal distribution have already been disputed in [14]) are different for every distribution F.

The approach of BENARD and BOS-LEVENBACH [2] may therefore be of interest. They aimed at median-unbiasedness instead of unbiasedness by choosing (cf. (3.2.14) and (3.2.15))

$$(3.2.19) \qquad P_i = F(m(\underline{x}_{i:n})) = \frac{i-\alpha'_{i:n}}{n+1-2\alpha'_{i:n}} .$$

The main advantage of this procedure is clearly that $\alpha'_{i:n}$ is independent of the distribution F, since $F(m(\underline{x}_{i:n})) = m(i,n)$. Endeavouring to find a single acceptable value α' for $\alpha'_{i:n}$ they computed

$$\lim_{n \to \infty} \alpha'_{i:n} = \alpha'_i$$

for fixed values of i, $1 \leq i \leq 100$. They found increasing values of α'_i from $\alpha'_1 = 0.307$ to $\alpha'_{100} = 0.333$. Noting this small variation it seemed useful to them to stress the sensitive extreme values and choose $\alpha' = 0.3$, or

$$P'_i = \frac{i - 0.3}{n + 0.4} \quad . \tag{3.2.20}$$

However, this reasoning seems to be incomplete. One might conjecture that

$$\lim_{i \to \infty} \alpha'_i = \lim_{i \to \infty} \lim_{n \to \infty} \alpha'_{i:n} \overset{?}{=} \lim_{r \to 0} \alpha'(r) = \lim_{r \to 0} \lim_{\substack{\frac{i}{n} \to r \\ n \to \infty}} \alpha'_{i:n} \quad ,$$

although it is well known that the asymptotic behaviour of extreme values (i fixed as n tends to infinity, or $r = 0$) is entirely different from the behaviour of the central order statistics ($0 < r < 1$) (cf. [4], 81 ff, where a similar conjecture for the $\alpha_{i:n}$ is shown to be correct for some classes of distributions). One might also easily guess from the numerical values found for α'_i, that $\lim_{i \to \infty} \alpha'_i = \frac{1}{3}$ +). Even then, as long as the behaviour of $\alpha'(r)$ for $0 < r < 1$ is entirely unknown, the choice of any particular compromise value α' seems hard to justify. It is thought that (3.2.16) completes a satisfactory justification of this interesting approach to the plotting problem.

+) Prof.Dr. J.Th. RUNNENBURG has kindly informed the author that while the paper by BENARD and BOS-LEVENBACH [2] was in preparation, Prof.Dr. H. KESTEN and he partly filled the gaps pointed out above. The result, however, went unpublished.

Chapter 4

TWO WEAK-ORDER RELATIONS FOR DISTRIBUTION FUNCTIONS

4.1. A WEAK ORDERING AND AN EQUIVALENCE FOR THE CLASS $\mathcal{F}$

In this chapter we return to the study of non-decreasing, convex and concave-convex transforms $\phi(\underline{x})$ of a random variable $\underline{x}$. As we did in chapter 3 we shall restrict ourselves to considering random variables $\underline{x}$, $\underline{x}^*$, $\underline{x}^{**}$,... having distribution functions F, F^*, $F^{**} \in \mathcal{F}$ (cf. (3.1.6) - (3.1.11)). The supports of these distributions will be denoted by I, I^*, I^{**},... , and their inverse functions by G, G^*, G^{**},... ; $\underline{x}_{i:n}$, $\underline{x}^*_{i:n}$,... denote the i-th order statistics of a sample of size n from the parent distributions F, F^*,... , and $F_{i:n}$, $F^*_{i:n}$,... their respective distribution functions.

Going over the theorems proved in chapter 2 we note that they are concerned exclusively with properties of the marginal distributions of $\underline{x}$ and $\phi(\underline{x})$, F and F^*, say, and not with their simultaneous distribution at all. It is therefore obvious that one may replace the random variable $\phi(\underline{x})$ by any other random variable $\underline{x}^*$ that is isomorous with $\phi(\underline{x})$ (i.e. has the same probability distribution F^* as $\phi(\underline{x})$) in any of these theorems. This amounts to the same thing as saying that these theorems are not concerned with the random variables $\underline{x}$ and $\phi(\underline{x})$ at all, but simply with the distributions F and F^*.

The following lemma is needed to bridge the gap between the approach of chapter 2 and the one outlined here.

LEMMA 4.1.1

For any pair $F, F^* \in \mathcal{F}$ there exists a strictly increasing function ϕ on I such that, if $\underline{x}$ has distribution function F then $\phi(\underline{x})$ has distribution function F^*. The function ϕ is uniquely defined on I by $\phi(x) = G^*F(x)$; it is twice differentiable on I with continuous second derivative ϕ''.

PROOF

It is necessary and sufficient that ϕ should satisfy

$$F^*\phi(x) = P(\phi(\underline{x}) \leq \phi(x)) = P(\underline{x} \leq x) = F(x) \quad ,$$

for all $x \in I$, and be strictly increasing on I. Obviously $\phi(x) = G^*F(x)$ is the only function that meets these requirements (cf. (3.1.6) and (3.1.7)). The last assertion of the lemma follows from (3.1.6) and (3.1.9).

Now we define the following order relation on $\mathcal{F}$:

DEFINITION 4.1.1

If $F, F^* \in \mathcal{F}$, then $F \underset{c}{<} F^*$ (or equivalently $F^* \underset{c}{>} F$) if and only if G^*F is convex on I.

We shall say in this case that F c-precedes F^* or that F^* c-follows F and that the two are c-comparable. We shall also speak of c-ordering, c-comparison, etc., where the letter c stands for convex. By lemma 4.1.1 the meaning of this definition is clear: $F \underset{c}{<} F^*$ if and only if a random variable with distribution F may be transformed into one with distribution F^* by an increasing, convex transformation. Furthermore we note that G^*F is (increasing) convex on I if and only if its inverse function GF^* is (increasing) concave on I^*. Hence concavity of GF^* on I^* might also have been used to define the relation $F \underset{c}{<} F^*$.

Obviously $F \underset{c}{<} F$ for all $F \in \mathcal{F}$; since an increasing, convex function of a convex function is again convex, $F \underset{c}{<} F^* \underset{c}{<} F^{**}$ yields $F \underset{c}{<} F^{**}$ for $F, F^*, F^{**} \in \mathcal{F}$. The relation $\underset{c}{<}$ is thus a weak ordering on $\mathcal{F}$. By defining an equivalence relation $\sim$ by

DEFINITION 4.1.2

If $F, F^* \in \mathcal{F}$, then $F \sim F^*$ if and only if $F \underset{c}{<} F^*$ and $F^* \underset{c}{<} F$,

and passing to the collection $\overline{\mathcal{F}}$ of equivalence classes we may define a partial ordering on $\overline{\mathcal{F}}$ by ordering equivalence classes according to the c-ordering of their representatives. The structure of the equivalence classes is given by

LEMMA 4.1.2

If $F, F^* \in \mathcal{F}$, then $F \sim F^*$ if and only if $F(x) = F^*(ax+b)$ for some constants $a > 0$ and b.

PROOF

$F \sim F^*$ if and only if G^*F is convex as well as concave on I and hence linear and increasing, or $G^*F(x) = ax+b$ or $F(x) = F^*(ax+b)$, $a > 0$.

In statistical parlance the lemma asserts that c-ordering is independent of location and scale parameters: the class $\overline{\mathcal{F}}$ is the class of types of laws belonging to $\mathcal{F}$. We may consequently restrict our attention to c-comparison of standardized distribution functions.

In what follows we shall often have to establish c-ordering of two specific distributions. A number of criteria for convexity of G^*F to be used then, are given in the following lemma.

LEMMA 4.1.3

If $F, F^* \in \mathcal{F}$, then $F \underset{c}{<} F^*$ if and only if one of the following conditions is satisfied:

(1) $\quad G^{*''}F(x)\, F'^{2}(x) + G^{*'}F(x)\, F''(x) \geqq 0$ for all $x \in I$;

(2) $\quad \dfrac{G^{*'}(y)}{G'(y)}$ is non-decreasing for $0 < y < 1$;

(3) $\quad \dfrac{G''(y)}{G'(y)} \leqq \dfrac{G^{*''}(y)}{G^{*'}(y)}$ for $0 < y < 1$;

(4) $\quad$ Condition (3) holds for all $y \in R$, where R is dense in $(0,1)$.

PROOF

Condition (1) is obtained by differentiating $\varphi(x) = G^{*}F(x)$ twice with respect to x and setting $\varphi''(x) \geqq 0$ on I. From $\phi(x) = G^{*}F(x)$ we find, by setting $x = G(y)$, $G^{*}(y) = \phi G(y)$ for $0 < y < 1$, and we have the definition: $F \underset{c}{<} F^{*}$ if and only if $G^{*}(y)$ is a convex function of (is convex in) $G(y)$ (cf. [13], 75). Differentiating $G^{*}(y)$ once or twice with respect to $G(y)$ we obtain conditions (2) and (3). Since both members of the inequality of condition (3) are continuous on $(0,1)$ by (3.1.9) and (3.1.10) condition (4) is also sufficient.

4.2. PROPERTIES OF c-ORDERED PAIRS OF DISTRIBUTIONS

Making use of the results of the previous chapters we may now easily obtain a number of theorems that state properties and provide characterizations of the order relation $\underset{c}{<}$ on $\mathcal{F}$. The first one is

THEOREM 4.2.1

Let R be a dense subset of $(0,1)$. Then for $F, F^{*} \in \mathcal{F}$ the following statements are equivalent:

(1) $\quad F \underset{c}{<} F^{*}$;

(2) $\quad F(E\, \underline{x}_{i:n}) \leqq F^{*}(E\, \underline{x}^{*}_{i:n})$ for all $n=1,2,\ldots$ and $i=1,2,\ldots,n$,

for which $E\,\underline{x}_{i:n}$ and $E\,\underline{x}^*_{i:n}$ exist;

(3) If i and n tend to infinity in such a way that $\lim \frac{i}{n} = r$, $r \in R$, then

$$\lim \; n\Big(F^*(E\,\underline{x}^*_{i:n}) - F(E\,\underline{x}_{i:n})\Big) \geqq 0 .$$

PROOF

By (3.1.1)

$$G^*_{i:n}F_{i:n}(x) = (B_{i:n}F^*)^{-1}B_{i:n}F(x) = G^*F(x) \quad , \tag{4.2.1}$$

and hence $F \underset{c}{<} F^*$ implies that $F_{i:n} \underset{c}{<} F^*_{i:n}$ since the latter distributions clearly belong to $\mathcal{F}$. Thus from statement (1) we find that $\underline{x}^*_{i:n}$ is isomorous with $\phi(\underline{x}_{i:n}) = G^*F(\underline{x}_{i:n})$, where ϕ is strictly increasing and convex on I. (This is also immediate from the fact that $\underline{x}^*$ is isomorous with $\phi(\underline{x})$, and that an increasing transformation does not disturb the ranks of the order statistics.)

Hence if $F \underset{c}{<} F^*$, by lemma 2.2.1,

$$G^*F(E\,\underline{x}_{i:n}) \leqq E\,G^*F(\underline{x}_{i:n}) = E\,\underline{x}^*_{i:n} \quad , \qquad \text{or}$$

$$F(E\,\underline{x}_{i:n}) \leqq F^*(E\,\underline{x}^*_{i:n}) \quad ,$$

which proves (1) $\Rightarrow$ (2). As (2) $\Rightarrow$ (3) is trivially true it remains to be proved that (3) yields (1).

By theorem 3.2.1

$$\lim \; n\Big(F^*(E\,\underline{x}^*_{i:n}) - F(E\,\underline{x}_{i:n})\Big) = \tfrac{1}{2}r(1-r)\left(\frac{G^{*''}(r)}{G^{*'}(r)} - \frac{G''(r)}{G'(r)}\right), \tag{4.2.2}$$

and consequently (3) implies that

$$\frac{G''(r)}{G'(r)} \leqq \frac{G^{*''}(r)}{G^{*'}(r)} \qquad \text{for all } r \in R \quad .$$

By lemma 4.1.3 we find $F \underset{c}{<} F^*$ which proves the theorem.

Theorem 4.2.1 presents two equivalent approaches to the problem of finding inequalities for expected values of order statistics by comparison with distributions for which these quantities are either analytically tractable or numerically known. The equivalence of (1) and (2) permits an approach by means of a convexity proof, whereas the equivalence of (2) and (3) enables one to start from known asymptotic results.

The second part of the theorem was originally proved by the present author giving a geometrically intuitive proof [27]. The regularity conditions imposed on F and F^* to ensure continuity of the second derivative of G^*F then acquire a natural significance. They ensure that, if G^*F would not be convex on I, it would necessarily be strictly concave on a sub-interval of I. Choosing r in this interval one shows that the inequality of statement (3) is false.

Finally we remark that as $F \underset{c}{<} F^*$ implies $F(E\,\underline{x}) \leqq F^*(E\,\underline{x}^*)$ we may expect distributions c-following on one another to show an increasing skewness to the right, c.q. a decreasing skewness to the left. Although a much stronger result in this direction is demonstrated in theorem 4.2.2 it is interesting to note that this application of JENSEN's inequality points this way too.

<u>THEOREM 4.2.2</u>

Let R be a dense subset of (0,1). Then for F, $F^* \in \mathcal{F}$ the following statements are equivalent:

(1) $F \underset{c}{<} F^*$;

(2) $\dfrac{\mu_{2k+1}(\underline{x}_{i:n})}{\sigma^{2k+1}(\underline{x}_{i:n})} \leqq \dfrac{\mu_{2k+1}(\underline{x}^*_{i:n})}{\sigma^{2k+1}(\underline{x}^*_{i:n})}$ for all k=1,2,... , n=1,2,...,

and i=1,2,...,n , for which $\mu_{2k+1}(\underline{x}_{i:n})$ and $\mu_{2k+1}(\underline{x}^*_{i:n})$ exist;

(3) If i and n tend to infinity in such a way that $\lim \frac{i}{n} = r$, $r \in R$, then for at least one value of k=1,2,...

$$\lim \sqrt{n}\left(\frac{\mu_{2k+1}(\underline{x}^{*}_{i:n})}{\sigma^{2k+1}(\underline{x}^{*}_{i:n})} - \frac{\mu_{2k+1}(\underline{x}_{i:n})}{\sigma^{2k+1}(\underline{x}_{i:n})}\right) \geqq 0$$

PROOF

If $F \underset{c}{\leq} F^{*}$, $\underline{x}^{*}_{i:n}$ is isomorous with $G^{*}F(\underline{x}_{i:n})$ where $G^{*}F$ is strictly increasing and convex on I (cf. the proof of theorem 4.2.1). Application of theorem 2.2.1 immediately yields (2). As (2) $\Rightarrow$ (3) trivially it remains to be proved that (3) implies (1).

By theorem 3.2.2 we have

$$(4.2.3) \qquad \lim \sqrt{n}\left(\frac{\mu_{2k+1}(\underline{x}^{*}_{i:n})}{\sigma^{2k+1}(\underline{x}^{*}_{i:n})} - \frac{\mu_{2k+1}(\underline{x}_{i:n})}{\sigma^{2k+1}(\underline{x}_{i:n})}\right) =$$

$$= 2^{-k}\,\frac{(2k+1)!}{(k-1)!}\,\sqrt{r(1-r)}\,\left(\frac{G^{*''}(r)}{G^{*'}(r)} - \frac{G''(r)}{G'(r)}\right) \quad ,$$

and hence we have from (3)

$$\frac{G''(r)}{G'(r)} \leqq \frac{G^{*''}(r)}{G^{*'}(r)} \qquad \text{for all } r \in R \quad .$$

By lemma 4.1.3 we find that $F \underset{c}{\leq} F^{*}$, which completes the proof.

We note that the large sample inequality (3) for one value of k implies the inequality (2) for finite sample size for all k.

We end this section with a theorem concerning the measure of skewness $\frac{E\,\underline{x} - m(\underline{x})}{\sigma(\underline{x})}$. In section 2.2 we have shown by a counter-example that this quantity does not necessarily increase by a non-decreasing, convex transform of the random variable $\underline{x}$. Theorem 4.2.3 shows that asymptotically for large sample order statistics from distributions in $\mathcal{F}$ no such anomalies can exist.

THEOREM 4.2.3

Let R be a dense subset of (0,1). Then for F, $F^{*} \varepsilon \mathcal{F}$ the following statements are equivalent:

(1) $F \underset{c}{<} F^{*}$;

(2) If i and n tend to infinity in such a way that $\lim \frac{i}{n} = r$, $0 < r < 1$, then

$$\lim \sqrt{n}\left(\frac{E\,\underline{x}^{*}_{i:n} - m(\underline{x}^{*}_{i:n})}{\sigma(\underline{x}^{*}_{i:n})} - \frac{E\,\underline{x}_{i:n} - m(\underline{x}_{i:n})}{\sigma(\underline{x}_{i:n})}\right) \geqq 0 \quad ;$$

(3) Statement (2) is valid for all $r \varepsilon R$.

PROOF

By theorem 3.2.3 statements (2) or (3) hold if and only if

$$\text{(4.2.4)} \qquad \frac{1}{2}\sqrt{r(1-r)}\left(\frac{G^{*''}(r)}{G^{*'}(r)} - \frac{G''(r)}{G'(r)}\right) \geqq 0$$

for all $0 < r < 1$ or all $r \varepsilon R$, and hence by lemma 4.1.3 if and only if $F \underset{c}{<} F^{*}$.

4.3. EXAMPLES OF c-ORDERING

In this section we shall discuss a number of examples of the order relation $\underset{c}{<}$ considered in the preceding sections. Especially the first three examples are meant to provide simple illustrations of the theory rather than sharp inequalities for use in specific cases.

4.3.1. c-COMPARISON WITH THE RECTANGULAR DISTRIBUTION

We take $F^*(x) = x$, $0 < x < 1$, or $G^*(y) = y$. Since

$$F^*(E\, \underline{x}^*_{i:n}) = \frac{i}{n+1} \quad \text{and} \quad \frac{\mu_3(\underline{x}^*_{i:n})}{\sigma^3(\underline{x}^*_{i:n})} = 2\,\frac{n+1-2i}{n+3}\sqrt{\frac{n+2}{i(n+1-i)}}$$

application of theorems 4.2.1 and 4.2.2 for $k = 1$ gives for $F \in \mathcal{F}$:

If the density function F' is non-decreasing on I (F convex), then

$$F(E\, \underline{x}_{i:n}) \leqq \frac{i}{n+1} \quad \text{and}$$

$$\frac{\mu_3(\underline{x}_{i:n})}{\sigma^3(\underline{x}_{i:n})} \leqq 2\,\frac{n+1-2i}{n+3}\sqrt{\frac{n+2}{i(n+1-i)}} \quad ;$$

if the density function is non-increasing on I (F concave) the inequalities are reversed. We note that the result for $F(E\, \underline{x}_{i:n})$ was mentioned by BLOM ([4], 68).

4.3.2. c-COMPARISON WITH $F^*(x) = -\frac{1}{x}$ AND $F^*(x) = \frac{x-1}{x}$

For $F^*(x) = -\frac{1}{x}$, $-\infty < x < -1$, or $G^*(y) = -\frac{1}{y}$ we find,

for $i \geqq 2$, $\quad F^*(E\, \underline{x}^*_{i:n}) = \frac{i-1}{n}$ and,

for $i \geqq 4$, $\quad \frac{\mu_3(\underline{x}^*_{i:n})}{\sigma^3(\underline{x}^*_{i:n})} = -2\,\frac{2n+1-i}{i-3}\sqrt{\frac{i-2}{n(n+1-i)}}$.

Application of theorem 4.2.1 and theorem 4.2.2 for $k = 1$ gives for $F \in \mathcal{F}$:

If $\frac{1}{F(x)}$ is concave on I then

$$F(E\ \underline{x}_{i:n}) \leqq \frac{i-1}{n} \qquad \text{for } i \geqq 2 \quad \text{and}$$

$$\frac{\mu_3(\underline{x}_{i:n})}{\sigma^3(\underline{x}_{i:n})} \leqq -2\,\frac{2n+1-i}{i-3}\sqrt{\frac{i-2}{n(n+1-i)}} \qquad \text{for } i \geqq 4 \quad ;$$

if $\frac{1}{F(x)}$ is convex on I the inequalities are reversed.

For $F^{*}(x) = \frac{x-1}{x}$, $1 < x < \infty$, or $G^{*}(y) = \frac{1}{1-y}$ we find

for $i \leqq n-1$, $\quad F^{*}(E\ \underline{x}^{*}_{i:n}) = \frac{i}{n}$ and,

for $i \leqq n-3$, $\quad \frac{\mu_3(\underline{x}^{*}_{i:n})}{\sigma^3(\underline{x}^{*}_{i:n})} = 2\,\frac{n+i}{n-i-2}\sqrt{\frac{n-i-1}{n\,i}}$,

and hence we have for $F \in \mathcal{F}$:

If $\frac{1}{1-F(x)}$ is convex on I then

$$F(E\ \underline{x}_{i:n}) \leqq \frac{i}{n} \qquad \text{for } i \leqq n-1 \quad \text{and}$$

$$\frac{\mu_3(\underline{x}_{i:n})}{\sigma^3(\underline{x}_{i:n})} \leqq 2\,\frac{n+i}{n-i-2}\sqrt{\frac{n-i-1}{n\,i}} \qquad \text{for } i \leqq n-3 \quad ;$$

if $\frac{1}{1-F(x)}$ is concave on I the inequalities are reversed.

Combining the results of paragraphs 4.3.1 and 4.3.2 we may set up crude bounds for the expected values of order statistics in terms of the distribution quantiles for many distribution functions, for instance

A. GAMMA DISTRIBUTIONS: $F'(x) = \frac{1}{\Gamma(\tau)}\, e^{-x}\, x^{\tau-1}$, $\tau > 0$, $0 < x < \infty$.

For $\tau \leqq 1$, F' is non-increasing and 4.3.1 is applicable. Furthermore one easily shows by repeated differentiation that $\frac{1}{F(x)}$

is convex for all values of τ, and $\frac{1}{1-F(x)}$ is convex for $\tau \geq 1$. Summarizing we obtain

$$\tau > 1 \qquad \frac{i-1}{n} \leq F(E\,\underline{x}_{i:n}) \leq \frac{i}{n}$$

$$\tau = 1 \qquad \left(\frac{i-1}{n} <\right) \frac{i}{n+1} \leq F(E\,\underline{x}_{i:n}) \leq \frac{i}{n}$$

$$\tau < 1 \qquad \left(\frac{i-1}{n} <\right) \frac{i}{n+1} \leq F(E\,\underline{x}_{i:n}) \;.$$

B. BETA DISTRIBUTIONS: $F'(x) = \frac{1}{B(\tau_1,\tau_2)} x^{\tau_1 - 1} (1-x)^{\tau_2 - 1}$, $\tau_1 > 0$, $\tau_2 > 0$, $0 < x < 1$.

F' is non-decreasing for $\tau_1 \geq 1$, $\tau_2 \leq 1$, and non-increasing for $\tau_1 \leq 1$, $\tau_2 \geq 1$. Repeated differentiation shows that $\frac{1}{F(x)}$ is convex for $\tau_2 \geq 1$ and $\frac{1}{1-F(x)}$ is convex for $\tau_1 \geq 1$. Hence we obtain

$$\tau_1 > 1,\ \tau_2 > 1 \qquad \frac{i-1}{n} \leq F(E\,\underline{x}_{i:n}) \leq \frac{i}{n}$$

$$\tau_1 > 1,\ \tau_2 = 1 \qquad \frac{i-1}{n} \leq F(E\,\underline{x}_{i:n}) \leq \frac{i}{n+1} \left(< \frac{i}{n}\right)$$

$$\tau_1 = 1,\ \tau_2 > 1 \qquad \left(\frac{i-1}{n} <\right) \frac{i}{n+1} \leq F(E\,\underline{x}_{i:n}) \leq \frac{i}{n}$$

$$\tau_1 \geq 1,\ \tau_2 < 1 \qquad F(E\,\underline{x}_{i:n}) \leq \frac{i}{n+1} \left(< \frac{i}{n}\right)$$

$$\tau_1 < 1,\ \tau_2 \geq 1 \qquad \left(\frac{i-1}{n} <\right) \frac{i}{n+1} \leq F(E\,\underline{x}_{i:n}) \;.$$

The case $\tau_1 = \tau_2 = 1$ is trivial and the case $\tau_1 < 1$, $\tau_2 < 1$ is not covered by either 4.3.1 or 4.3.2.

C. NORMAL DISTRIBUTION: $F'(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$, $-\infty < x < \infty$.

Here $\frac{1}{F(x)}$ and $\frac{1}{1-F(x)}$ are both convex, so we find

$$\frac{i-1}{n} \leq F(E\,\underline{x}_{i:n}) \leq \frac{i}{n} \quad ,$$

corresponding to $\tau_1 \to \infty$ and $\tau_1, \tau_2 \to \infty$ in cases A and B, as is proper. We note that in all these cases the bounds $\frac{i-1}{n}$ and $\frac{i}{n}$ derived from 4.3.2 hold trivially for $i = 1$ and $i = n$ respectively. We need hardly point out that inequalities of the same type may be given for the skewness of order statistics from these distributions.

4.3.3. c-COMPARISON WITH THE EXPONENTIAL DISTRIBUTION

c-Comparison with the exponential distribution was recently discussed in an entirely different context by R.E. BARLOW, A.W. MARSHALL and F. PROSCHAN [1]. For a continuous distribution F let the hazard rate q be defined by

$$q(x) = \frac{F'(x)}{1-F(x)} \quad ;$$

for a life distribution this is the conditional probability density of failure at time x, given performance up to time x. In [1] the authors studied distributions having monotone hazard rate, which are of practical interest in reliability theory.

The distribution F is said to have increasing hazard rate if q is non-decreasing on I, and decreasing hazard rate if q is non-increasing on I. From these definitions it is easily seen that if F has decreasing hazard rate, then I is necessarily of the form $I = (a,\infty)$; if F has monotone hazard rate this implies that $F'(x) > 0$ on I (cf. (3.1.7)).

Let us restrict ourselves to distributions $F \in \mathcal{F}$ and let $F^*(x) = 1 - e^{-x}$, $0 < x < \infty$, be the exponential distribution. Setting

$$\phi(x) = G^* F(x) = -\log\,(1 - F(x)) \quad ,$$

we have $\phi'(x) = q(x)$ on I and hence:

F has increasing hazard rate if and only if $F \underset{c}{<} F^*$;
F has decreasing hazard rate if and only if $F \underset{c}{>} F^*$.

The study of distributions with monotone hazard rate is therefore a study of distributions that are c-comparable with the exponential distribution. The main results of BARLOW, MARSHALL and PROSCHAN in [1] are that the class of distributions with increasing hazard rate is closed under convolution, and that the class of distributions with decreasing hazard rate and the same support $I = (a,\infty)$ is closed under convex combination. We remark that the class $\mathcal{F}$ is also closed under convex combination of distributions with the same support. Hence, denoting the convolution of F_1 and F_2 by $F_1 * F_2$, we have in the terminology of the present study:

Let F_1, F_2, $F_1 * F_2 \in \mathcal{F}$ and let F^* denote the exponential distribution. If $F_1 \underset{c}{<} F^*$ and $F_2 \underset{c}{<} F^*$ then $F_1 * F_2 \underset{c}{<} F^*$; if $F_1 \underset{c}{>} F^*$ and $F_2 \underset{c}{>} F^*$, and F_1 and F_2 have the same support $I_1 = I_2$, then $\lambda F_1 + (1-\lambda)F_2 \underset{c}{>} F^*$ for $0 \leqq \lambda \leqq 1$.

The expected values of order statistics from the exponential distribution are easy to find:

$$E\,\underline{x}^*_{i:n} = \sum_{j=0}^{i-1} \frac{1}{n-j} \leqq \int_{n-i+\frac{1}{2}}^{n+\frac{1}{2}} \frac{1}{x}\,dx = \log \frac{n+\frac{1}{2}}{n-i+\frac{1}{2}} ,$$

and hence

$$F^*(E\,\underline{x}^*_{i:n}) \leqq \frac{i}{n+\frac{1}{2}} .$$

We note that this inequality cannot be sharpened for all i and n, since from (3.2.10) we find $\alpha(r) = 0$, $\beta(r) = \frac{1}{2}$, and hence the inequality is asymptotically sharp as $\frac{i}{n}$ tends to r, for all $0 < r < 1$. From theorem 4.2.1 we now find for instance:

If $F \in \mathcal{F}$ has increasing hazard rate ($F \underset{c}{<} F^*$) then

$$F(E\,\underline{x}_{i:n}) \leqq \frac{i}{n+\frac{1}{2}} .$$

4.3.4. THE MAXIMAL c-CHAIN OF GAMMA DISTRIBUTIONS

In a partially ordered set we define a chain to be a subset of which any two elements are comparable. A maximal chain is a chain which is not a proper subset of any other chain. We recall KURATOWSKI's lemma stating that any partially ordered set contains at least one maximal chain (cf. [26]).

If we start looking for a c-chain in the partially ordered class $\overline{\mathcal{F}}$ of types of laws belonging to $\mathcal{F}$ (cf. section 4.1) and keep in mind that c-ordering implies increasing skewness, a plausible candidate seems to be the class of gamma distributions

$$F_\tau(x) = \frac{1}{\Gamma(\tau)} \int_0^x e^{-t}\, t^{\tau-1}\, dt \quad , \quad 0 < x < \infty, \quad \tau > 0 .$$

We shall first sketch a proof that $F_{\tau'} \underset{c}{<} F_\tau$ for $0 < \tau < \tau'$, i.e. the gamma distributions c-follow one another with decreasing values of the parameter.

This means we have to prove that

$$\phi(x) = G_{\tau'} F_\tau(x) \quad , \qquad 0 < \tau < \tau' ,$$

is concave for $0 < x < \infty$, where $G_{\tau'}$ denotes the inverse of $F_{\tau'}$. The fact that no explicit expression for $G_{\tau'}$ is available leads to the following indirect approach. Consider the function

$$\psi(x) = F_\tau(x) - F_{\tau'}(b(x+a)) \quad ,$$

for $b > 0$, $a \geqq 0$ and $0 \leqq x \leqq \infty$. As $F_\tau(x) - F_{\tau'}\phi(x) \equiv 0$, and $F_{\tau'}$ is strictly increasing, $\psi(x)$ has the same sign as $\phi(x) - b(x+a)$ for all $x \geqq 0$. Also

$$\psi'(x) = F_\tau'(x) - b\, F_{\tau'}'(b(x+a))$$

has the same sign as

$$\chi(x) = \log F_\tau'(x) - \log F_{\tau'}'(b(x+a)) - \log b \quad ,$$

and we have

$$\chi'(x) = (b-1) + \frac{\tau-1}{x} - \frac{\tau'-1}{x+a}$$

A detailed and laborious study of the sign of $\chi'(x)$ for $x \geqq 0$ and various sets of values of a, b, τ and τ', and of the signs of $\chi(x)$ and $\psi(x)$ for $x = 0$ and $x \to \infty$ reveals that $\psi(x)$, and hence $\phi(x) - b(x+a)$, can have at most two distinct zeros and is positive between these zeros. For $b > 0$, $a < 0$ a comparison with the case $b > 0$, $a = 0$ shows that $\phi(x) - b(x+a)$ can have at most one zero, whereas for $b \leqq 0$ the same conclusion holds since ϕ is strictly increasing. Thus the graph of ϕ lies above every chord which proves concavity of ϕ.

To construct a maximal c-chain we add the normal distribution F_∞ and the class of distributions

$$F_{-\tau}(x) = 1 - F_\tau(-x) \quad , \qquad \tau > 0, \quad -\infty < x < 0,$$

to the family of gamma distributions F_τ , $\tau > 0$. Now

$$G_{-\tau'}F_{-\tau}(x) = G_{-\tau'}(1 - F_\tau(-x)) = -\, G_{\tau'}F_\tau(-x)$$

is convex for $0 < \tau < \tau'$, $x < 0$, so $F_{-\tau} \underset{c}{<} F_{-\tau'}$ for $0 < \tau < \tau'$. Also $F_\infty \underset{c}{<} F_\tau$ for all $\tau > 0$ since $G_\infty F_\tau$ is the limit of the (standardized) concave functions $G_{\tau'}F_\tau$, $0 < \tau < \tau'$, $\tau' \to \infty$; $F_{-\tau} \underset{c}{<} F_\infty$ for all $\tau > 0$ follows by the same argument. Hence the class F_τ , $-\infty < \tau \leqq +\infty, \tau \neq 0$, is indeed a c-chain in $\overline{\mathcal{F}}$, where

$$F_\tau \underset{c}{<} F_{\tau'} \qquad \text{for } \frac{1}{\tau} \leqq \frac{1}{\tau'} \quad .$$

To show that the c-chain is maximal we remark that $F_\tau \underset{c}{<} F \underset{c}{<} F_{\tau'}$ for fixed τ and all $\frac{1}{\tau'} > \frac{1}{\tau}$, implies that $GF_\tau(x) = \lim_{\tau' \to \tau} GF_{\tau'}(x)$ is convex as well as concave, and hence that F and F_τ are equivalent and may be identified. Similarly one shows that $F_\tau \underset{c}{<} F \underset{c}{<} F_{\tau'}$ for fixed τ' and all $\frac{1}{\tau} < \frac{1}{\tau'}$, implies that $F \sim F_{\tau'}$. It remains to be proved that no element of $\overline{\mathcal{F}}$ can be added at either end of the c-chain.

For $\tau > 0$, $x > 0$,

$$1 - F_\tau(x) = \frac{1}{\Gamma(\tau)} \int_x^\infty e^{-t}\, t^{\tau-1}\, dt \geqq \frac{1}{\Gamma(\tau)} \int_x^{2x} e^{-t}\, t^{\tau-1}\, dt \geqq$$

$$\geqq \frac{e^{-2x} x^\tau}{\Gamma(\tau)} \int_x^{2x} \frac{dt}{t} = \log 2 \; \frac{e^{-2x} x^\tau}{\Gamma(\tau)} \quad .$$

Denoting, for the purpose of this proof only, by $\underline{x}_{i:n}(\tau)$ the order statistic $\underline{x}_{i:n}$ from the distribution F_τ we have from (3.1.4)

$$E\, \underline{x}_{1:n}(\tau) \geqq \frac{n(\log 2)^{n-1}}{(\Gamma(\tau))^n} \int_0^\infty e^{-(2n-1)x}\, x^{n\tau}\, dx =$$

$$= \frac{n(\log 2)^{n-1}(2n-1)^{-n\tau-1}}{(\Gamma(\tau))^n}\, \Gamma(n\tau+1) \geqq c_n\, \tau^n$$

for τ sufficiently small, where c_n is a positive constant independent of τ. Hence

$$0 \leqq 1 - F_\tau(E\, \underline{x}_{i:n}(\tau)) \leqq 1 - F_\tau(E\, \underline{x}_{1:n}(\tau)) \leqq \frac{1}{\Gamma(\tau)} \int_{c_n \tau^n}^\infty e^{-t}\, t^{\tau-1}\, dt$$

where the right-hand side may easily be shown to vanish as $\tau \to 0$, and as a result

$$\lim_{\tau \downarrow 0} F_\tau(E\, \underline{x}_{i:n}(\tau)) = 1$$

for all i and n.

Now for any distribution $F \varepsilon \mathcal{F}$ for which $F_\tau \underset{c}{<} F$ for all τ, we should have $F(E\, \underline{x}_{i:n}) = 1$ for all i and n by theorem 4.2.1. This would imply that either F is degenerate, or $E\, \underline{x}_{i:n}$ does not exist (as a finite quantity) for any i and n. Both possibilities contradict the assumption that $F \varepsilon \mathcal{F}$ by (3.1.6) and (3.1.8). In an analogous way one shows that no $F \varepsilon \mathcal{F}$ exists with $F \underset{c}{<} F_\tau$ for all τ. This concludes the proof that the c-chain is maximal.

To illustrate the results obtained in this section table 4.3.4.1 shows the values of $F(E\, \underline{x}_{i:10})$, i=1,2,...,10, for the gamma distributions F_τ , τ=1,2,...,5, and the normal distribution F_∞. For gamma distributions up to $\tau = 5$ values of $E\, \underline{x}_{i:n}$ are given by S.S. GUPTA [11], whereas the expected values of normal order statistics were taken from D. TEICHROEW [24].

TABLE 4.3.4.1

Values of $F(E\, \underline{x}_{i:10})$ for gamma distributions with parameter τ .

	τ=1	τ=2	τ=3	τ=4	τ=5	$\tau=\infty$
i=1	0.095	0.080	0.075	0.072	0.071	0.062
2	0.190	0.177	0.172	0.170	0.168	0.158
3	0.285	0.274	0.269	0.267	0.266	0.256
4	0.381	0.370	0.367	0.365	0.363	0.354
5	0.476	0.467	0.464	0.462	0.461	0.451
6	0.571	0.563	0.560	0.559	0.558	0.549
7	0.666	0.660	0.657	0.656	0.655	0.646
8	0.760	0.756	0.754	0.753	0.752	0.744
9	0.855	0.851	0.850	0.849	0.848	0.842
10	0.947	0.945	0.944	0.943	0.943	0.938

We note that the inequalities derived in paragraph 4.3.2 are indeed rather crude. On the other hand the smooth appearance of curves of the tabled values for fixed i suggests that computation of $E\, \underline{x}_{i:n}$ for different values of τ may perhaps largely proceed by interpolation for $F(E\, \underline{x}_{i:n})$ with respect to τ, for which the monotonicity proved in this paragraph provides a firm basis.

Since for $\tau = 1$ we have the explicit expression $G_1(y) = -\log(1-y)$ available, a part of the above result is easy to obtain, viz. that F_τ c-precedes or c-follows the exponential distribution F_1 for $\tau \geqq 1$ or $\tau \leqq 1$. BARLOW, MARSHALL and PROSCHAN [1] used this as an example of their theorem that the convolution of distributions c-preceding F_1 again c-precedes F_1, and as a counterexample that the convolution of distributions c-following F_1 does not necessarily c-follow F_1 (cf. paragraph 4.3.3). Also from paragraph 4.3.3 we find that for gamma distributions with $\tau \geqq 1$

$$F(E\, \underline{x}_{i:n}) \leqq \frac{i}{n+\frac{1}{2}} \quad .$$

We have noted before that for $\tau = 1$ this inequality is asymptotically sharp; table 4.3.4.1 shows that it is fairly sharp already for moderate sample size.

4.4. A WEAK ORDERING FOR A CLASS OF SYMMETRIC DISTRIBUTIONS

In the remaining part of this chapter we restrict our attention to the subclass $\mathcal{S} \subset \mathcal{F}$ of symmetric distributions F defined by (cf. section 2.1)

(4.4.1) $$F \in \mathcal{F}$$

(4.4.2) $$F(x_0-x) + F(x_0+x) = 1 \qquad \text{for some } x_0 \text{ and all real } x.$$

Condition (4.4.2) may also be written

$$(4.4.3) \qquad G(y) + G(1-y) = 2x_0 \qquad \text{for all } 0 < y < 1 ,$$

and hence for $F \in \mathcal{S}$ and $F^* \in \mathcal{S}$

$$G^*F(x_0-x) + G^*F(x_0+x) = 2x_0^* \quad ,$$

where x_0^* denotes the point of symmetry of F^*. This means that G^*F is antisymmetrical on I about $x_0 \in I$, and consequently convexity (c.q. concavity) of G^*F for $x > x_0$, $x \in I$, implies concavity (c.q. convexity) of G^*F for $x < x_0$, $x \in I$, and conversely. It follows immediately that if $F, F^* \in \mathcal{S}$, then $F \underset{c}{<} F^*$ implies $F \sim F^*$, i.e. symmetric distributions are not c-comparable unless they are equivalent.

We may, however, define a different order relation on $\mathcal{S}$ which is better adapted to the situation:

DEFINITION 4.4.1

If $F, F^* \in \mathcal{S}$, then $F \underset{s}{<} F^*$ (or equivalently $F^* \underset{s}{>} F$) if and only if G^*F is convex for $x > x_0$, $x \in I$.

We shall say in this case that F s-precedes F^* or that F^* s-follows F and that the two are s-comparable. We shall also speak of s-ordering, s-comparison, etc., where the letter s stands for symmetry. By lemma 4.1.1 we see that $F \underset{s}{<} F^*$ means that a random variable with distribution F may be transformed into one with distribution F^* by an increasing, antisymmetrical, concave-convex transformation. We note that G^*F is concave-convex on I if and only if GF^* is convex-concave on I^*.

Clearly $F \underset{s}{<} F$ for all $F \in \mathcal{S}$; since G^*F maps x_0 on x_0^* and an increasing, convex function of a convex function is again convex, $F \underset{s}{<} F^* \underset{s}{<} F^{**}$ yields $F \underset{s}{<} F^{**}$ for $F, F^*, F^{**} \in \mathcal{S}$. The relation $\underset{s}{<}$ is thus a weak ordering on $\mathcal{S}$. Defining an equivalence relation $\underset{s}{\sim}$ by

DEFINITION 4.4.2

If $F, F^* \in \mathcal{S}$, then $F \underset{s}{\sim} F^*$ if and only if $F \underset{s}{<} F^*$ and $F^* \underset{s}{<} F$,

and passing to the collection $\overline{\mathcal{S}}$ of equivalence classes, the relation $\underset{s}{<}$ defines a partial ordering on $\overline{\mathcal{S}}$. Analogous to lemma 4.1.2 we find

LEMMA 4.4.1

If $F, F^* \in \mathcal{S}$, then $F \underset{s}{\sim} F^*$ if and only if $F(x) = F^*(ax+b)$ for some constants $a > 0$ and b.

PROOF

$F \underset{s}{\sim} F^*$ if and only if G^*F is concave-convex as well as convex-concave on I and hence linear and increasing, or $G^*F(x) = ax+b$, or $F(x) = F^*(ax+b)$, $a > 0$.

The lemma shows that this order relation is also independent of location and scale parameters. The symbol $\underset{s}{\sim}$ is superfluous and may be replaced by $\sim$.

We round off this section by giving the analogue of lemma 4.1.3:

LEMMA 4.4.2

If $F, F^* \in \mathcal{S}$, then $F \underset{s}{<} F^*$ if and only if one of the following conditions is satisfied:

(1) $G^{*''}F(x)\, F'^2(x) + G^{*'}F(x)\, F''(x) \geqq 0$ for all $x > x_0$, $x \in I$;

(2) $\dfrac{G^{*'}(y)}{G'(y)}$ is non-decreasing for $\frac{1}{2} < y < 1$;

(3) $\dfrac{G''(y)}{G'(y)} \leqq \dfrac{G^{*''}(y)}{G^{*'}(y)}$ for $\frac{1}{2} < y < 1$;

(4) Condition (3) holds for all $y \in R$ where R is dense in $(\frac{1}{2},1)$.

PROOF

The proof is analogous to that of lemma 4.1.3.

4.5. PROPERTIES OF s-ORDERED PAIRS OF DISTRIBUTIONS

In this section the properties of s-ordering will be seen to resemble closely the properties of c-ordering discussed in section 4.2. The first theorem provides a necessary and sufficient condition for s-ordering in terms of inequalities for $F(E\,\underline{x}_{i:n})$.

THEOREM 4.5.1

Let R be a dense subset of $(\frac{1}{2},1)$. Then for $F, F^{*} \in \mathcal{S}$ the following statements are equivalent:

(1) $F \underset{s}{<} F^{*}$;

(2) $F(E\,\underline{x}_{i:n}) \leqq F^{*}(E\,\underline{x}^{*}_{i:n})$ for all $n=1,2,\ldots$ and $\frac{n+1}{2} \leqq i \leqq n$, for which $E\,\underline{x}^{*}_{i:n}$ exists;

(3) If i and n tend to infinity in such a way that $\lim \frac{i}{n} = r$, $r \in R$, then

$$\lim\; n\Big(F^{*}(E\,\underline{x}^{*}_{i:n}) - F(E\,\underline{x}_{i:n})\Big) \;\geqq\; 0$$

PROOF

For $i \geqq \frac{n+1}{2}$ one easily finds from (3.1.1) and (4.4.2) that the probability density $F'_{i:n}$ of $\underline{x}_{i:n}$ satisfies

$$F'_{i:n}(x_0+x) - F'_{i:n}(x_0-x) \;\geqq\; 0 \qquad \text{for } x \geqq 0 ,$$

and hence

$$F_{i:n}(x_0+x) + F_{i:n}(x_0-x)$$

is a non-decreasing function of x for $x \geqq 0$. Here x_0 denotes the point of symmetry of F (cf. (4.4.2)).

Furthermore, $\underline{x}^*_{i:n}$ is isomorous with $\phi(\underline{x}_{i:n}) = G^*F(\underline{x}_{i:n})$ and if $F \underset{s}{<} F^*$, ϕ is increasing, antisymmetrical, concave-convex on I about x_0. Hence existence of E $\underline{x}^*_{i:n}$ obviously implies existence of E $\underline{x}_{i:n}$, and the conditions of theorem 2.3.1 are satisfied for the random variable $\underline{x}_{i:n}$ and the function $\phi(x) = G^*F(x)$. Upon application of the theorem we obtain for $i \geqq \frac{n+1}{2}$

$$G^*F(E\ \underline{x}_{i:n}) \leqq E\ G^*F(\underline{x}_{i:n}) = E\ \underline{x}^*_{i:n} \quad ,$$

or

$$F(E\ \underline{x}_{i:n}) \leqq F^*(E\ \underline{x}^*_{i:n}) \quad ,$$

which proves that (1) gives (2). Appealing to lemma 4.4.2 instead of lemma 4.1.3 the remainder of the proof follows that of theorem 4.2.1 verbatim.

We note that again, as with theorem 4.2.1, the large sample result (3) is equivalent to statement (2) for finite sample size. We also remark that since $F(E\ \underline{x}_{i:n}) \leqq F^*(E\ \underline{x}^*_{i:n})$ implies $F_{i:n}(E\ \underline{x}_{i:n}) \leqq F^*_{i:n}(E\ \underline{x}^*_{i:n})$ the theorem indicates that for $i \geqq \frac{n+1}{2}$ the order statistics of distributions s-following on one another will have a tendency towards increasing skewness to the right c.q. decreasing skewness to the left. The following theorem is concerned with this point.

THEOREM 4.5.2

Let R be a dense subset of $(\frac{1}{2},1)$. Then for F, $F^* \in \mathcal{S}$ the following statements are equivalent:

(1) $\quad F \underset{s}{<} F^*$;

(2) $\quad$ If i and n tend to infinity in such a way that $\lim \frac{i}{n} = r$, $\frac{1}{2} < r < 1$, then for all k=1,2,...

$$\lim \sqrt{n} \left(\frac{\mu_{2k+1}(\underline{x}^*_{i:n})}{\sigma^{2k+1}(\underline{x}^*_{i:n})} - \frac{\mu_{2k+1}(\underline{x}_{i:n})}{\sigma^{2k+1}(\underline{x}_{i:n})} \right) \geqq 0 \quad ;$$

(3) Statement (2) is valid for all $r \in R$ and at least one value of $k=1,2,\ldots$.

PROOF

By (4.2.3) statements (2) and (3) are equivalent to

$$\frac{G''(r)}{G'(r)} \leqq \frac{G^{*''}(r)}{G^{*'}(r)}$$

for all $\frac{1}{2} < r < 1$ and all $r \in R$ respectively. By lemma 4.4.2 both are equivalent to $F \underset{s}{\leq} F^*$.

We note that the inequality (3) for one value of k implies the inequality (2) for all k. The author has not been able to prove or disprove a corresponding small sample inequality.

Finally we have a counterpart of theorem 4.2.3:

THEOREM 4.5.3

Let R be a dense subset of $(\frac{1}{2},1)$. Then for F, $F^* \in \mathcal{S}$ the following statements are equivalent:

(1) $F \underset{s}{\leq} F^*$;

(2) If i and n tend to infinity in such a way that $\lim \frac{i}{n} = r$, $\frac{1}{2} < r < 1$, then

$$\lim \sqrt{n} \left(\frac{E\,\underline{x}^*_{i:n} - m(\underline{x}^*_{i:n})}{\sigma(\underline{x}^*_{i:n})} - \frac{E\,\underline{x}_{i:n} - m(\underline{x}_{i:n})}{\sigma(\underline{x}_{i:n})} \right) \geqq 0 \quad ;$$

(3) Statement (2) is valid for all $r \in R$.

PROOF

Appealing to lemma 4.4.2 instead of lemma 4.1.5 the proof closely resembles that of theorem 4.2.3.

We conclude this section by remarking that for $i \leqq \frac{n+1}{2}$, $0 < r < \frac{1}{2}$ and R dense in $(0,\frac{1}{2})$, the inequalities of theorems 4.5.1, 4.5.2 and 4.5.3 are reversed.

4.6. EXAMPLES OF s-ORDERING

The first two examples given in this section are similar to those treated in section 4.3. They refer to s-comparison with the rectangular distribution and to mutual s-comparison of symmetric beta distributions. The third example treats the s-ordering of the normal and logistic distributions.

4.6.1. s-COMPARISON WITH THE RECTANGULAR DISTRIBUTION

We take $F^*(x) = x$, $0 < x < 1$, or $G^*(y) = y$, and $F^*(E\,\underline{x}^*_{i:n}) = \frac{i}{n+1}$. For $F \in \mathcal{S}$ we consider symmetric distributions having density functions F' that possess a single extreme, and are therefore either U-shaped (single minimum; F concave-convex) or unimodal (single maximum; F convex-concave). By theorem 4.5.1 we have for $F \in \mathcal{S}$:

For a symmetric, U-shaped distribution, $F(E\,\underline{x}_{i:n}) \leqq \frac{i}{n+1}$ for $i \geqq \frac{n+1}{2}$;
for a symmetric, unimodal distribution, $F(E\,\underline{x}_{i:n}) \geqq \frac{i}{n+1}$ for $i \geqq \frac{n+1}{2}$.

BLOM ([4], 66) proved the latter inequality asymptotically for large samples, which by theorem 4.5.1 is equivalent to the result stated here. D. VAN DANTZIG and J. HEMELRIJK [9] mention the result for all n in connection with a comparison of TERRY's and VAN DER WAERDEN's tests where respectively the quantities $E\,\underline{x}_{i:n}$ and $G(\frac{i}{n+1})$ for the normal distribution are involved in the test statistic.

4.6.2. THE s-CHAIN OF SYMMETRIC BETA DISTRIBUTIONS

Consider the class of distribution functions

$$F_\tau(x) = \frac{\Gamma(2\tau)}{2^{2\tau-1}(\Gamma(\tau))^2} \int_{-1}^{x} (1-t^2)^{\tau-1}\, dt \quad ,$$

$\tau > 0$, $-1 < x < 1$, that are equivalent to the symmetric beta distributions (the random variable has been transformed by $\underline{x} = 2\underline{u} - 1$). We shall sketch a proof that $F_\tau \underset{s}{<} F_{\tau'}$ for $0 < \tau < \tau'$, i.e. the symmetric beta distributions s-follow one another with increasing values of the parameter.

Hence we have to show that

$$\phi(x) = G_{\tau'} F_\tau(x) \quad , \qquad 0 < \tau < \tau' \quad ,$$

is convex for $0 < x < 1$, where $G_{\tau'}$ denotes the inverse of $F_{\tau'}$. As in paragraph 4.3.4 we consider the function

$$\psi(x) = F_\tau(x) - F_{\tau'}(b(x+a)) \quad ,$$

for $b > 0$, $ba \geqq -1$, $b(1+a) \leqq 1$, which has the same sign as $\phi(x) - b(x+a)$ for $0 \leqq x \leqq 1$. Also $\psi'(x)$ has the same sign as

$$\chi(x) = \log F'_\tau(x) - \log F'_{\tau'}(b(x+a)) - \log b \quad ,$$

and we have

$$\chi'(x) = -\frac{2(\tau-1)x}{1-x^2} + \frac{2b^2(\tau'-1)(x+a)}{1-b^2(x+a)^2} \quad .$$

As in 4.3.4 we study the sign of $\chi'(x)$ for $0 \leqq x \leqq 1$ and the signs of

$\chi(x)$ and $\psi(x)$ for $x = 0$ and $x = 1$. In this way we find that $\psi(x)$ and hence $\phi(x) - b(x+a)$ can have at most two zeros for $0 \leqq x \leqq 1$, in which case the function is negative between these zeros. For $b > 0$, $ba < -1$, $b(1+a) \leqq 1$, the representation of $F_\tau(x)$ and $F_{\tau'}(b(x+a))$ by the beta integrals remains valid for $-a-\frac{1}{b} \leqq x \leqq 1$ and hence we may prove the same result in this interval. However, $\phi(x) - b(x+a) < 0$ for $0 \leqq x \leqq -a$, and hence the result continues to apply for $0 \leqq x \leqq 1$. For $b > 0$, $b(1+a) > 1$ a comparison with the case $b > 0$, $b(1+a) = 1$ shows that $\phi(x) - b(x+a)$ can have at most one zero for $0 \leqq x \leqq 1$; for $b \leqq 0$ the same holds since ϕ is strictly increasing. Hence for $0 \leqq x \leqq 1$ the graph of ϕ lies below any chord which proves convexity of ϕ for $0 \leqq x \leqq 1$.

4.6.3. s-COMPARISON OF NORMAL AND LOGISTIC DISTRIBUTIONS

Consider

$$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}t^2} dt , \qquad -\infty < x < \infty,$$

and

$$F^*(x) = \frac{1}{1+e^{-x}} , \qquad -\infty < x < \infty.$$

Clearly $F, F^* \in \mathcal{I}$; furthermore one easily shows by repeated differentiation that

$$G^* F(x) = \log F(x) - \log (1-F(x)) ,$$

is convex for $x \geqq 0$, so $F \underset{s}{\leq} F^*$.

Now $E\, \underline{x}^*_{i:n}$ is simple to evaluate, giving for $i \geqq \frac{n+1}{2}$,

$$E\, \underline{x}^*_{i:n} = \sum_{k=n+1-i}^{i-1} \frac{1}{k} \leqq \log \frac{i-\frac{1}{2}}{n-i+\frac{1}{2}} ,$$

and hence

$$F^*(E\,\underline{x}^*_{i:n}) \leqq \frac{i-\frac{1}{2}}{n} \quad .$$

Applying theorem 4.5.1 we find

$$F(E\,\underline{x}_{i:n}) \leqq \frac{i-\frac{1}{2}}{n} \qquad \text{for} \quad i \geqq \frac{n+1}{2} \quad .$$

We note that BLOM [4] proved the corresponding asymptotic result for $n \to \infty$. The inequality cannot be sharpened for all $i \geqq \frac{n+1}{2}$ and n since $\alpha(1) = \lim_{r \to 1} \alpha(r) = \frac{1}{2}$ in (3.2.12) for the normal distribution.

The easy derivation of this inequality, and of those obtained in paragraphs 4.3.1, 4.3.2, 4.3.3 and 4.6.1, is a consequence of the fact that in all these cases G^* is an incomplete beta function. The properties of these distributions make them particularly well suited as standards for c-comparison and s-comparison. The symmetrical distributions of this type will be studied in the next chapter.

4.7. GENERALIZATION TO OTHER DISTRIBUTIONS

In chapters 3 and 4 we have confined our research to the class $\mathcal{F}$ of distribution functions satisfying (3.1.6) - (3.1.8) or (3.1.9) - (3.1.11). It may be of interest to discuss briefly the extent to which our results remain valid for other types of distributions. We shall consider the relation $\underset{c}{<}$; the conclusions for s-ordering are similar.

First we remark that condition (3.1.8) (or (3.1.11)) seems hard to avoid. Without this, few non-trivial results are to be obtained.

Let us therefore consider the class $\mathcal{F}'$ of non-degenerate distributions F satisfying (3.1.8). Though these distribution functions do not necessarily possess unique inverse functions we may define $F \underset{c}{<} F^*$, for $F, F^* \in \mathcal{F}'$, if and only if there exists a function ϕ which is convex (and non-decreasing) on I and

$$F^{*}\phi(x) = F(x) \qquad \text{on } I \ .$$

Obviously, with this definition we still find that $\underline{x}^{*}$ is isomorous with $\phi(\underline{x})$ (cf. section 4.1), and the results of chapter 2 are applicable. Hence in theorems 4.2.1 and 4.2.2 the result that statement (1) ($F \underset{c}{<} F^{*}$) yields statement (2) (the small sample inequality) remains valid. The same holds for theorem 4.5.1 concerning s-ordering.

However, the remaining part of the results of this chapter rapidly break down as soon as condition (3.1.7) is seriously violated. Let $\mathcal{F}''$ be the class of continuous, but not necessarily strictly increasing, distribution functions F satisfying (3.1.6) and (3.1.8), where the relation $\underset{c}{<}$ may be defined as above. The inverse function G is defined and twice continuously differentiable up to discontinuities and we note that $F \underset{c}{<} F^{*}$ implies that G and G^{*} have the same points of discontinuity. From the proofs given it is clear that for $F \in \mathcal{F}''$ the large sample results of section 3.2 continue to hold if $r = \lim \frac{i}{n}$ is a continuity point of G. However, if $\lim \frac{i}{n} = r$, $0 < r < 1$, and G is discontinuous at $y = r$, then for sufficiently large n one easily shows

$$G(r-0) \leqq E\, \underline{x}_{i:n} = \int_0^1 G(y)\, b_{i:n}(y)\, dy \leqq G(r+0) \ ,$$

and consequently

$$F(E\, \underline{x}_{i:n}) = r \ .$$

Hence, if both G and G^{*} are discontinuous at $y = r$, $0 < r < 1$, we have

$$\lim_{\frac{i}{n} \to r} n\left(F^{*}(E\, \underline{x}^{*}_{i:n}) - F(E\, \underline{x}_{i:n})\right) = 0 \ .$$

This affords a simple counter-example to the second part of theorem 4.2.1, which shows that for $F, F^{*} \in \mathcal{F}''$ the large sample inequalities of the theorem do not imply c-comparability of F and F^*. Suppose $F(x) = p$, $0 < p < 1$, for $a \leqq x \leqq b$ whereas it is strictly increasing for $x < a$ or $x > b$ on I. Let the function ϕ be defined on $x < a$, $x \in I$, and $x > b$, $x \in I$, let it be strictly increasing, twice continuously differentiable and convex on both these intervals, with $\phi(a-0) < \phi(b+0)$, and let $E\,\phi(\underline{x}_{i:n})$ exist for some i and n. The distribution F^* of the random variable $\underline{x}^*$ that is isomorous with $\phi(\underline{x})$ obviously belongs to $\mathcal{F}''$ and both G and G^* have only one discontinuity at $y = p$. Furthermore, if $F^*\psi(x) = F(x)$ on I, then $\psi(x) = \phi(x)$ for $x < a$ and $x > b$, $x \in I$. This means, however, that unless

$$\phi'(a-0) \leqq \frac{\phi(b+0) - \phi(a-0)}{b-a} \leqq \phi'(b+0) \quad ,$$

ψ can not possibly be convex on I. On the other hand, we have shown above that if $\lim \frac{i}{n} = r$, then

$$\lim \quad n\left(F^*(E\,\underline{x}^*_{i:n}) - F(E\,\underline{x}_{i:n})\right) \geqq 0$$

for all $0 < r < 1$, which establishes the counter-example.

Finally one might discard condition (3.1.6) and consider e.g. discrete distributions. Results in this direction do not, however, appear to be of great interest since a non-decreasing, convex transformation of a discrete random variable simply changes the values it assumes while leaving the corresponding probabilities unaffected (of course some probabilities in the extreme left tail may be added if ϕ remains constant on some sub-interval of I). Thus one cannot c-compare (or s-compare) essentially different distributions as one may in the continuous case. In fact, hardly anything seems to be gained by thinking in terms of distributions at all, instead of in terms of random variables as we did in chapter 2.

Chapter 5

s-COMPARISON WITH SYMMETRIC INVERSE BETA DISTRIBUTIONS

5.1. SYMMETRIC INVERSE BETA DISTRIBUTIONS

In this chapter we shall confine our attention to the class $\mathcal{S}$ of symmetric distributions defined by (4.4.1) and (4.4.2). For $F, F^* \in \mathcal{S}$ we found in theorem 4.5.1 that it is possible to obtain inequalities $F(E\,\underline{x}_{i:n}) \leqq F^*(E\,\underline{x}^*_{i:n})$ for all n and $i \geqq \frac{n+1}{2}$ from the corresponding large sample inequalities or, equivalently, from the relation $F \underset{s}{\leq} F^*$. Writing (cf. (3.2.11) and (3.2.12))

$$F(E\,\underline{x}_{i:n}) = \frac{i-\alpha_{i:n}}{n+1-2\alpha_{i:n}} \quad , \quad \alpha_{i:n} = \alpha_{n-i+1:n} \quad , \tag{5.1.1}$$

$$F^*(E\,\underline{x}^*_{i:n}) = \frac{i-\alpha^*_{i:n}}{n+1-2\alpha^*_{i:n}} \quad , \quad \alpha^*_{i:n} = \alpha^*_{n-i+1:n} \quad ,$$

and for $n \to \infty$, $\frac{i}{n} \to r$, $\frac{1}{2} < r < 1$,

$$\alpha(r) = \alpha(1-r) = \lim_{\frac{i}{n} \to r} \alpha_{i:n} = \frac{r(1-r)}{2(2r-1)} \frac{G''(r)}{G'(r)} \quad , \tag{5.1.2}$$

$$\alpha^*(r) = \alpha^*(1-r) = \lim_{\frac{i}{n} \to r} \alpha^*_{i:n} = \frac{r(1-r)}{2(2r-1)} \frac{G^{*''}(r)}{G^{*'}(r)} \quad ,$$

we note that $\frac{i-\alpha}{n+1-2\alpha}$ is a non-decreasing function of α for $i \geqq \frac{n+1}{2}$,

$\alpha \leqq \frac{n+1}{2}$, and hence that $F \underset{s}{<} F^*$ or $\alpha(r) \leqq \alpha^*(r)$ for all $\frac{1}{2} < r < 1$ implies $\alpha_{i:n} \leqq \alpha^*_{i:n}$ for all i and n.

It seems to be interesting to investigate the possibility to obtain inequalities for $\alpha_{i:n}$ for specific distributions F that hold for all i and n by s-comparison with a member F^* of some class of distribution functions that are comparatively easy to handle. In the first place this class has to be such that E $\underline{x}^*_{i:n}$ is sufficiently tractable to obtain inequalities for $\alpha^*_{i:n}$ for all i and n. Secondly, if e.g. one wishes to determine an upper bound for $\alpha_{i:n}$, it will have to be fairly simple to establish the order relation $F \underset{s}{<} F^*$, or $\alpha(r) \leqq \alpha^*(r)$ for $\frac{1}{2} < r < 1$, which in general can be rather complicated (cf. paragraph 4.6.2). One may in part overcome the latter difficulty by choosing F^* in such a way that

$$\sup_{\frac{1}{2} < r < 1} \alpha(r) \leqq \inf_{\frac{1}{2} < r < 1} \alpha^*(r) ,$$

which in most cases should be easy to accomplish. The resulting inequalities for $\alpha_{i:n}$, however, will not in general be asymptotically sharp. An obvious exception is of course the case where $\alpha^*(r)$ does not depend on r, or

$$\text{(5.1.3)} \qquad \frac{y(1-y)}{2(2y-1)} \frac{G^{*''}(y)}{G^{*'}(y)} = \text{constant} \quad \text{for } \frac{1}{2} < y < 1 .$$

We are thus led to consider as standards for s-comparison the class of distribution functions F^*_τ satisfying (5.1.3) or

$$\text{(5.1.4)} \qquad G^{*'}_\tau(y) = a\, y^{-\tau} (1-y)^{-\tau} , \qquad -\infty < \tau < \infty .$$

Since our methods are independent of location and scale parameters we may, for the sake of definiteness, define

(5.1.5) $$G_\tau^*(y) = \int_{\frac{1}{2}}^{y} u^{-\tau} (1-u)^{-\tau}\, du \quad , \qquad -\infty < \tau < \infty .$$

F_τ^* will be called the symmetric inverse beta distribution with parameter τ. It is easy to see that $F_\tau^* \in \mathcal{S}$ and that its density is unimodal for $\tau > 0$ and U-shaped for $\tau < 0$. We note that this class contains three well-known distributions; for $\tau = 0, \frac{1}{2}$, and 1 we have

$$F_0^*(x) = x \quad , \qquad 0 < x < 1 ,$$

$$F_{\frac{1}{2}}^*(x) = \sin^2 x \quad , \qquad 0 < x < \frac{\pi}{2} ,$$

$$F_1^*(x) = \frac{1}{1+e^{-x}} \quad , \qquad -\infty < x < \infty ,$$

being the uniform, sine and logistic distributions. From

(5.1.6) $$\frac{G_\tau^{*''}(y)}{G_\tau^{*'}(y)} = \tau \frac{2y-1}{y(1-y)}$$

and (5.1.2) we find that for F_τ^* , $\alpha^*(r) = \frac{\tau}{2}$ for all $\frac{1}{2} < r < 1$. For F_τ^* , $\underline{x}_{i:n}^*$ and $\alpha_{i:n}^*$ will be denoted by $\underline{x}_{i:n}^*(\tau)$ and $\alpha_{i:n}^*(\tau)$. We note that $E\, \underline{x}_{i:n}^*(\tau)$ exists if $\tau-1 < i < n+2-\tau$.

Happily enough these distributions also meet our first requirement that $E\, \underline{x}_{i:n}^*(\tau)$ be to some extent manageable. By partial integration one obtains

(5.1.7) $$E\, \underline{x}_{k+1:n}^*(\tau) = E\, \underline{x}_{k:n}^*(\tau) + \frac{\Gamma(n+1)\ \Gamma(k+1-\tau)\ \Gamma(n-k+1-\tau)}{\Gamma(n+2-2\tau)\ \Gamma(k+1)\ \Gamma(n-k+1)}$$

provided both expectations exist. Furthermore, if for $i \geq \frac{n+1}{2}$ $E\, \underline{x}_{i:n}^*(\tau)$ exists, then $E\, \underline{x}_{k:n}^*(\tau)$ exists for $\frac{n}{2} \leq k \leq i$ and

$$(5.1.8)\qquad \begin{aligned} &E\,\underline{x}^{*}_{\frac{n+1}{2}:n}(\tau) = 0 && \text{if } n \text{ is odd},\\ &E\,\underline{x}^{*}_{\frac{n}{2}:n}(\tau) + E\,\underline{x}^{*}_{\frac{n}{2}+1:n}(\tau) = 0 && \text{if } n \text{ is even},\end{aligned}$$

since F^{*}_{τ} is symmetrical about $x = 0$, and as a result one may express $E\,\underline{x}^{*}_{i:n}(\tau)$ as a finite sum.

The properties of F^{*}_{τ} given above were previously discussed by G. BLOM in [4], who also considered asymmetrical inverse beta distributions. Among these are the exponential distribution and the distributions discussed in paragraph 4.3.2. As we remarked at the end of section 4.6 the simple behaviour of the inverse beta distributions explains the ease with which the examples concerning them were dealt with in chapter 4.

Returning to the symmetric distributions F^{*}_{τ}, it is obvious from lemma 4.4.2 and (5.1.6) that these distributions form an s-chain; they s-follow one another with increasing values of the parameter τ. Since it is usually fairly easy to determine what members of the s-chain s-follow or s-precede a given distribution $F \in \mathcal{S}$, the only thing that remains to be done is to find inequalities for $\alpha^{*}_{i:n}(\tau)$. After deriving some preliminary results in the next section we shall return to this problem in section 5.3.

5.2. INEQUALITIES FOR GAMMA AND BETA FUNCTIONS

The results established in this section will be needed in the sequel. Some of them may, however, be of independent interest.

We start by studying the quantities $a_n(x)$ defined for $n = 0,1,2,\ldots$, $-n < x < \infty$ by

$$n + a_n(x) > 0 ,$$

(5.2.1) $$\frac{\Gamma(n+1)}{\Gamma(n+x)} = (n + a_n(x))^{1-x} \qquad \text{for } x \neq 1 ,$$

$$a_n(1) = \lim_{x \to 1} a_n(x) = \exp\left(\frac{\Gamma'(n+1)}{\Gamma(n+1)}\right) - n .$$

Clearly $a_n(x)$ is a continuous function of x for $x > -n$. Furthermore we observe that the function $\log \Gamma(x)$ is convex for $x > 0$, since

$$\Gamma(x) = \int_0^\infty e^{-t}\, t^{x-1}\, dt ,$$

$$\Gamma'(x) = \int_0^\infty \log t\; e^{-t}\, t^{x-1}\, dt ,$$

$$\Gamma''(x) = \int_0^\infty (\log t)^2\, e^{-t}\, t^{x-1}\, dt ,$$

and hence $\Gamma''(x)\Gamma(x) > \Gamma'^2(x)$ by SCHWARZ's inequality.

We shall prove five lemmata that provide some insight in the behaviour of $a_n(x)$. The first one of these is concerned with an upper bound for $a_n(x)$ for $0 \leqq x \leqq 2$.

LEMMA 5.2.1

For $n = 0,1,2,\dots$,

$$a_n(x) \leqq x \qquad \text{for } 0 \leqq x \leqq 1,\ n+x \neq 0 .$$

$$a_n(x) \leqq 1 \qquad \text{for } 1 \leqq x \leqq 2 .$$

PROOF

Since $\log \Gamma(x)$ is convex we have for $0 \leqq x \leqq 1$, $n+x \neq 0$,

$$\log \Gamma(n+1) = \log \Gamma\Big(x(n+x) + (1-x)(n+x+1)\Big) \leqq$$

$$\leqq x \log \Gamma(n+x) + (1-x) \log \Gamma(n+x+1) ,$$

or $\log \frac{\Gamma(n+1)}{\Gamma(n+x)} \leqq (1-x) \log (n+x)$, or $a_n(x) \leqq x$.

For $1 \leqq x \leqq 2$,

$$\log \Gamma(n+x) = \log \Gamma\Big((2-x)(n+1) + (x-1)(n+2)\Big) \leqq$$
$$\leqq (2-x) \log \Gamma(n+1) + (x-1) \log \Gamma(n+2) \quad ,$$

or $\log \frac{\Gamma(n+x)}{\Gamma(n+1)} \leqq (x-1) \log (n+1)$, or $a_n(x) \leqq 1$.

In a similar way one may derive lower bounds for $a_n(x)$ for $0 \leqq x \leqq 2$, and also bounds for different values of x.

The next three lemmata are concerned with the sequence $a_n(x)$ for a fixed value of x. Here it will be tacitly understood that n runs through all non-negative integers satisfying $n > -x$, since otherwise $a_n(x)$ is not defined by (5.2.1).

LEMMA 5.2.2

For any fixed value of x, $\lim_{n \to \infty} a_n(x) = \frac{x}{2}$.

PROOF

For $x \neq 1$ we have by STIRLING's formula for $n \to \infty$

$$\log \frac{\Gamma(n+1)}{\Gamma(n+x)} = (n+\tfrac{1}{2}) \log (n+1) + (n-\tfrac{1}{2}+x) \log (n+x) - (1-x) +$$
$$+ \frac{1}{12}\left(\frac{1}{n+1} - \frac{1}{n+x}\right) + \mathcal{O}(n^{-3}) =$$
$$= (1-x)\left[\log n + \frac{x}{2n} + \frac{x(1-2x)}{12n^2}\right] + \mathcal{O}(n^{-3}) \quad ,$$

and hence

$$n + a_n(x) = \left(\frac{\Gamma(n+1)}{\Gamma(n+x)}\right)^{\frac{1}{1-x}} = n.\exp\left[\frac{x}{2n} + \frac{x(1-2x)}{12n^2} + \mathcal{O}(n^{-3})\right] =$$

$$= n + \frac{x}{2} + \frac{x(2-x)}{24\,n} + \mathcal{O}(n^{-2}) \quad ,$$

or

$$(5.2.2) \qquad a_n(x) = \frac{x}{2} + \frac{x(2-x)}{24n} + \mathcal{O}(n^{-2}) \quad .$$

For $x = 1$ the result of the lemma follows by continuity. We note that for sufficiently large n the sequence $a_n(x)$ is decreasing for $0 < x < 2$ and increasing for $x < 0$ or $x > 2$.

LEMMA 5.2.3

For a fixed value of x, the sequence $a_n(x)$ is convex if $0 \leqq x \leqq 2$ and concave if $x \leqq 0$, $x \geqq 2$, i.e.

$$2a_{n+1}(x) \leqq a_n(x) + a_{n+2}(x) \qquad \text{for } 0 \leqq x \leqq 2 \quad ,$$

$$2a_{n+1}(x) \geqq a_n(x) + a_{n+2}(x) \qquad \text{for } x \leqq 0 \quad \text{or } x \geqq 2 \quad .$$

PROOF

For $0 \leqq x \leqq 2$, $x \neq 1$, consider the function

$$f_n(x) = \left(\frac{n+x}{n+1}\right)^{\frac{1}{1-x}} + \left(\frac{n+2}{n+1+x}\right)^{\frac{1}{1-x}} - 2 \quad .$$

Obviously $\lim_{n \to \infty} f_n(x) = 0$. To show that $f_n(x) \geqq 0$ it is therefore sufficient to show that (treating n as a continuous variable)

$$\frac{df_n(x)}{dn} = \left(\frac{n+x}{n+1}\right)^{\frac{x}{1-x}} (n+1)^{-2} - \left(\frac{n+2}{n+1+x}\right)^{\frac{x}{1-x}} (n+1+x)^{-2} \leqq 0 \;,$$

or that

$$\frac{x}{1-x}\log(n+x) - \frac{2-x}{1-x}\log(n+1) \leqq \frac{x}{1-x}\log(n+2) - \frac{2-x}{1-x}\log(n+1+x) \quad ,$$

or that

$$g_n(x) = \frac{1}{1-x}\left[x \log\frac{n+2}{n+x} - (2-x)\log\frac{n+1+x}{n+1}\right] \geqq 0 \quad .$$

The last statement is certainly true since $\lim_{n\to\infty} g_n(x) = 0$ and

$$\frac{dg_n(x)}{dn} = \frac{1}{1-x}\left[\frac{x}{n+2} - \frac{x}{n+x} + \frac{2-x}{n+1} - \frac{2-x}{n+1+x}\right] =$$

$$= -\frac{x(2-x)}{(n+1)(n+2)(n+x)(n+1+x)} \leqq 0 \quad ,$$

and hence we have proved that $f_n(x) \geqq 0$ for $0 \leqq x \leqq 2$, $x \neq 1$. For $x \leqq 0$ or $x \geqq 2$ we may repeat the above proof while reversing all inequalities to show that $f_n(x) \leqq 0$.

Now

$$\left(\frac{n+x}{n+1}\right)^{\frac{1}{1-x}} = \frac{n+a_n(x)}{n+1+a_{n+1}(x)} \quad ,$$

$$\left(\frac{n+2}{n+1+x}\right)^{\frac{1}{1-x}} = \frac{n+2+a_{n+2}(x)}{n+1+a_{n+1}(x)} \quad ,$$

and hence

$$f_n(x) = \frac{a_n(x)+a_{n+2}(x)-2a_{n+1}(x)}{n+1+a_{n+1}(x)} \quad .$$

Since the denominator is positive by (5.2.1) the lemma is proved for $x \neq 1$; for $x = 1$ it follows by continuity.

We remark that, of course, the lemma implies that for $0 \leqq x \leqq 2$ and integers $k < n < l$

$$(5.2.3)\qquad a_n(x) \leqq \frac{1-n}{1-k}\, a_k(x) + \frac{n-k}{1-k}\, a_1(x) \quad ,$$

whereas for $x \leqq 0$ or $x \geqq 2$ the inequality is reversed.

LEMMA 5.2.4

For a fixed value of x, the sequence $a_n(x)$ is non-increasing if $0 \leqq x \leqq 2$ and non-decreasing if $x \leqq 0$ or $x \geqq 2$.

PROOF

As a convex (concave) sequence having a finite limit is non-increasing (non-decreasing) the lemma follows from lemmata 5.2.2 and 5.2.3.

For $x \neq 0$, $x \neq 2$, this monotonicity of $a_n(x)$ is asymptotically strict as may be seen from (5.2.2); by the convexity (concavity) it is therefore strict for all n. For $x = 0$ and $x = 2$ we have $a_n(0) = 0$, $a_n(2) = 1$ for all n.

From lemma 5.2.4 one finds the following inequalities which may have some interest of their own. For fixed integer k and all integer $n \geqq k$ one has

$$(5.2.4)\qquad \left(n+\frac{x}{2}\right)^{1-x} \leqq \frac{\Gamma(n+1)}{\Gamma(n+x)} \leqq \left(n+a_k(x)\right)^{1-x} \quad \text{for } 0 \leqq x \leqq 1 \text{ or } x \geqq 2;$$

for $1 \leqq x \leqq 2$ or $x \leqq 0$, the inequalities are reversed. In table 5.2.1 some values for $a_n(x)$ are given for $0 \leqq x \leqq 2$. The rapid decrease of $a_n(x)$ with n indicates that the inequalities (5.2.4) will be very sharp even for moderate values of k.

Having considered $a_n(x)$ for fixed x as a sequence with index n we might also discuss its behaviour for fixed n as a function of x for $x > -n$. It is easy to show that $a_n(x)$ is an increasing function of x. We shall not, however, need this in the sequel where the following lemma will suffice.

TABLE 5.2.1

Values of $a_n(x)$ for $0 \leqq x \leqq 2$

x \ n	0	1	2	3	5	10	20	∞
0	-	0	0	0	0	0	0	0
0.1	0.08184	0.05698	0.05377	0.05257	0.05156	0.05079	0.05039	0.05
0.2	0.14881	0.11262	0.10697	0.10478	0.10293	0.10148	0.10075	0.1
0.3	0.20900	0.16712	0.15964	0.15666	0.15411	0.15209	0.15105	0.15
0.4	0.26506	0.22062	0.21181	0.20823	0.20510	0.20261	0.20132	0.2
0.5	0.31831	0.27324	0.26354	0.25949	0.25592	0.25304	0.25154	0.25
0.6	0.36951	0.32509	0.31483	0.31047	0.30657	0.30339	0.30172	0.3
0.7	0.41914	0.37625	0.36573	0.36117	0.35705	0.35366	0.35186	0.35
0.8	0.46753	0.42678	0.41626	0.41162	0.40737	0.40384	0.40196	0.4
0.9	0.51491	0.47675	0.46644	0.46181	0.45753	0.45394	0.45202	0.45
1	0.56146	0.52621	0.51629	0.51176	0.50754	0.50396	0.50203	0.5
1.1	0.60730	0.57519	0.56582	0.56148	0.55740	0.55390	0.55201	0.55
1.2	0.65255	0.62375	0.61506	0.61099	0.60711	0.60377	0.60194	0.6
1.3	0.69727	0.67191	0.66402	0.66028	0.65668	0.65356	0.65184	0.65
1.4	0.74154	0.71970	0.71272	0.70936	0.70611	0.70327	0.70169	0.7
1.5	0.78540	0.76715	0.76117	0.75825	0.75541	0.75290	0.75151	0.75
1.6	0.82890	0.81428	0.80937	0.80695	0.80458	0.80247	0.80128	0.8
1.7	0.87209	0.86111	0.85735	0.85547	0.85362	0.85196	0.85102	0.85
1.8	0.91498	0.90766	0.90510	0.90382	0.90253	0.90138	0.90072	0.9
1.9	0.95761	0.95395	0.95265	0.95199	0.95133	0.95072	0.95038	0.95
2	1	1	1	1	1	1	1	1

LEMMA 5.2.5

For $n+x > 0$,

$$a_{2n}(2x) \leqq 2a_n(x) \qquad \text{if } x \geqq 0$$

$$a_{2n}(2x) \geqq 2a_n(x) \qquad \text{if } x \leqq 0 .$$

PROOF

Consider the function

$$f_n(x) = x\left[\log\left(n+a_n(x)\right) - \log\left(n+\frac{1}{2}a_{2n}(2x)\right)\right] =$$

$$= \frac{x}{1-x}\log\frac{\Gamma(n+1)}{\Gamma(n+x)} - \frac{x}{1-2x}\log\frac{\Gamma(2n+1)}{\Gamma(2n+2x)} + x\log 2 ,$$

defined for $x = \frac{1}{2}$ and $x = 1$ by continuity. We shall prove the lemma by showing that $f_n(x) \geqq 0$ for $n+x > 0$.

As $\lim_{n\to\infty} f_n(x) = 0$ it is sufficient to show that the function

$$\Delta f_n(x) = f_{n+1}(x) - f_n(x) = \frac{x}{1-x}\log\frac{n+1}{n+x} - \frac{x}{1-2x}\log\frac{(2n+1)(2n+2)}{(2n+2x)(2n+2x+1)} =$$

$$= x\left(\frac{1}{1-x} - \frac{1}{1-2x}\right)\log\frac{n+1}{n+x} - \frac{x}{1-2x}\log\frac{2n+1}{2n+2x+1} =$$

$$= -\frac{x^2}{1-2x}\left[\frac{1}{1-x}\log\frac{n+1}{n+x} + \frac{1}{x}\log\frac{2n+1}{2n+2x+1}\right]$$

is non-positive. Treating n as a continuous variable we have

$$\frac{d\Delta f_n(x)}{dn} = \frac{x^2}{1-2x}\left[\frac{1}{(n+x)(n+1)} - \frac{4}{(2n+1)(2n+2x+1)}\right] =$$

$$= \frac{x^2}{(n+x)(n+1)(2n+1)(2n+2x+1)} \geqq 0 ;$$

as $\lim_{n\to\infty} \Delta f_n(x) = 0$, it follows that $\Delta f_n(x) \leqq 0$ which completes the proof.

Having established these properties of the functions $a_n(x)$ we proceed to apply these results to obtain inequalities for a ratio of beta functions. For all real τ and integer k and n satisfying

$$(5.2.5) \qquad \begin{array}{c} 1 \leq k \leq n-1 \\ \tau-1 < k < n+1-\tau \end{array}$$

we introduce the abbreviation

$$(5.2.6) \qquad P_{k,n}(\tau) = \frac{\Gamma(n+1)\ \Gamma(k+1-\tau)\ \Gamma(n-k+1-\tau)}{\Gamma(n+2-2\tau)\ \Gamma(k+1)\ \Gamma(n-k+1)}$$

and prove

LEMMA 5.2.6

For integer k and n satisfying (5.2.5)

$$P_{k,n}(\tau) \geq \frac{(n+1-\tau)^{2\tau-1}}{\left(k+\frac{1}{2}-\frac{\tau}{2}\right)^{\tau}\left(n-k+\frac{1}{2}-\frac{\tau}{2}\right)^{\tau}} \qquad \text{if } -1 \leq \tau \leq 0,$$

$$P_{k,n}(\tau) \leq \frac{(n+1-\tau)^{2\tau-1}}{\left(k+\frac{1}{2}-\frac{\tau}{2}\right)^{\tau}\left(n-k+\frac{1}{2}-\frac{\tau}{2}\right)^{\tau}} \qquad \text{if } 0 \leq \tau \leq 1,$$

$$P_{k,n}(\tau) \leq \frac{n(n+2-2\tau)}{k(n-k+1-\tau)} \frac{(n+1-\tau)^{2\tau-3}}{\left(k-\frac{\tau}{2}\right)^{\tau-1}\left(n-k+1-\frac{\tau}{2}\right)^{\tau-1}} \qquad \text{if } 1 \leq \tau \leq 2,$$

and

$$P_{k,n}(\tau) \geq \frac{n(n+2-2\tau)}{k(n-k+1-\tau)} \frac{(n+1-\tau)^{2\tau-3}}{\left(k-\frac{\tau}{2}\right)^{\tau-1}\left(n-k+1-\frac{\tau}{2}\right)^{\tau-1}} \qquad \text{if } \tau \geq 2 .$$

PROOF

As both members of the first two inequalities are symmetrical in k and (n-k) it is sufficient to prove this part of the lemma for $k \geqq \frac{n}{2}$. By (5.2.1) we have

$$(5.2.7) \qquad P_{k,n}(\tau) = \frac{\left(n + a_n(2-2\tau)\right)^{2\tau-1}}{\left(k + a_k(1-\tau)\right)^{\tau}\left(n-k + a_{n-k}(1-\tau)\right)^{\tau}} .$$

Suppose first that $-1 \leqq \tau \leqq 0$. As $1 \leqq 1-\tau \leqq 2$ and $2 \leqq 2-2\tau \leqq 4$ we have from lemmata 5.2.2 and 5.2.4

$$a_k(1-\tau) \geqq \frac{1-\tau}{2} , \qquad a_{n-k}(1-\tau) \geqq \frac{1-\tau}{2} \qquad \text{and} \qquad a_n(2-2\tau) \leqq 1-\tau .$$

Since all exponents in (5.2.7) are non-positive this implies the result of the lemma for $-1 \leqq \tau \leqq 0$.

To prove the lemma for $0 \leqq \tau \leqq 1$ we start once more from (5.2.7) and apply lemma 5.2.5 for $x = 1-\tau \geqq 0$ to both factors of the denominator to obtain

$$P_{k,n}(\tau) \leqq \frac{\left(n + a_n(2-2\tau)\right)^{2\tau-1}}{\left(k + \frac{1}{2}a_{2k}(2-2\tau)\right)^{\tau}\left(n-k + \frac{1}{2}a_{2n-2k}(2-2\tau)\right)^{\tau}} .$$

By (5.2.3)

$$a_n(2-2\tau) \leqq \frac{1}{2}a_{2n-2k}(2-2\tau) + \frac{1}{2}a_{2k}(2-2\tau) \quad ;$$

furthermore $-\frac{1}{2}na - \frac{1}{4}a^2$ is a decreasing function of a for $n+a > 0$, thus for $a_n(2-2\tau) \leqq a \leqq \frac{1}{2}a_{2n-2k}(2-2\tau) + \frac{1}{2}a_{2k}(2-2\tau)$ by (5.2.1). Hence (omitting the argument (2-2τ) and writing a_n , a_{2n-2k} and a_{2k})

$$(k + \tfrac{1}{2}a_{2k})(n-k + \tfrac{1}{2}a_{2n-2k}) = (k + \tfrac{1}{2}a_n)(n-k + \tfrac{1}{2}a_n) - \tfrac{1}{2}na_n - \tfrac{1}{4}a_n^2 +$$

$$+ \tfrac{1}{2}(n-k)a_{2k} + \tfrac{1}{2}ka_{2n-2k} + \tfrac{1}{4}a_{2k}a_{2n-2k} \geqq$$

$$\geqq \ (k+\tfrac{1}{2}a_n)(n-k+\tfrac{1}{2}a_n) + \tfrac{1}{4}(2k-n)(a_{2n-2k}-a_{2k}) - \tfrac{1}{16}(a_{2n-2k}+a_{2k})^2 + $$
$$+ \tfrac{1}{4}a_{2k}a_{2n-2k} =$$

$$= \ (k+\tfrac{1}{2}a_n)(n-k+\tfrac{1}{2}a_n) + \tfrac{1}{16}(a_{2n-2k}-a_{2k})(8k-4n-a_{2n-2k}+a_{2k}) .$$

For $k = \frac{n}{2}$ the second term in the last member is equal to zero. For $k \geqq \frac{n+1}{2}$ it is non-negative since $a_{2n-2k} - a_{2k} \geqq 0$ by lemma 5.2.4, $a_{2k} \geqq 0$ by lemmata 5.2.2 and 5.2.4, and hence $a_{2n-2k} - a_{2k} \leqq a_{2n-2k} \leqq 1$ by lemma 5.2.1. Therefore

$$(k+\tfrac{1}{2}a_{2k})(n-k+\tfrac{1}{2}a_{2n-2k}) \geqq (k+\tfrac{1}{2}a_n)(n-k+\tfrac{1}{2}a_n) ,$$

and

$$(5.2.8) \qquad P_{k,n}(\tau) \leqq \frac{\left(n+a_n(2-2\tau)\right)^{2\tau-1}}{\left(k+\frac{1}{2}a_n(2-2\tau)\right)^{\tau}\left(n-k+\frac{1}{2}a_n(2-2\tau)\right)^{\tau}} \quad \text{for } 0 \leqq \tau \leqq 1.$$

As the right-hand side of (5.2.8) is obviously a decreasing function of $a_n(2-2\tau)$ for $\tau \geqq 0$, application of lemmata 5.2.2 and 5.2.4 yields

$$(5.2.9) \qquad P_{k,n}(\tau) \leqq \frac{(n+1-\tau)^{2\tau-1}}{\left(k+\frac{1}{2}-\frac{\tau}{2}\right)^{\tau}\left(n-k+\frac{1}{2}-\frac{\tau}{2}\right)^{\tau}} \quad \text{for } 0 \leqq \tau \leqq 1,$$

which proves the lemma for $0 \leqq \tau \leqq 1$.

Two more remarks should, however, be made about the preceding part of the proof. The first one is simply that for $0 \leqq \tau < 1$ we may extend definition (5.2.6) of $P_{k,n}(\tau)$ to the case where $k = 0$ or $k = n$. From the proof given it is clear that inequality (5.2.9) will continue to hold in this case. Secondly we note that all inequalities arrived at for $0 \leqq \tau \leqq 1$ in the above are reversed for $\tau \geqq 1$ and hence that

$$\text{(5.2.10)} \qquad P_{k,n}(\tau) \geqq \frac{(n+1-\tau)^{2\tau-1}}{\left(k+\frac{1}{2}-\frac{\tau}{2}\right)^{\tau}\left(n-k+\frac{1}{2}-\frac{\tau}{2}\right)^{\tau}} \qquad \text{for } \tau \geqq 1 ,$$

the only additional change in the proof being that for $k \geqq \frac{n}{2}$ $(8k - 4n - a_{2n-2k} + a_{2k})$ is now non-negative simply because $a_{2k} - a_{2n-2k} \geqq 0$.

Suppose now that $\tau \geqq 1$ and write

$$P_{k,n}(\tau) = \frac{n(n+2-2\tau)}{k(n-k+1-\tau)} \; P_{k-1,n-1}(\tau-1) \quad ,$$

where, according to (5.2.5), the latter quantity is defined if for $0 \leqq \tau-1 < 1$ we extend the definition of $P_{k-1,n-1}(\tau-1)$ to the case where $k-1 = 0$ (cf. the first remark following (5.2.9)). Application of (5.2.9) and (5.2.10) to the right-hand member yields at once the result of the lemma for $\tau \leqq 2$ and $\tau \geqq 2$ respectively.

We next prove

<u>LEMMA 5.2.7</u>

For integer k and n satisfying (5.2.5) and $k \geqq \frac{n}{2}$

$$\left(\frac{(k-\frac{\tau}{2})(n-k+1-\frac{\tau}{2})}{(k+1-\frac{\tau}{2})(n-k-\frac{\tau}{2})}\right)^{\tau-1} \geqq 1 + (\tau-1)\,\frac{2k-n}{k(n-k+1-\tau)} \qquad \text{if } 1 < \tau \leqq 2;$$

the inequality is reversed if $\tau \geqq 2$.

<u>PROOF</u>

For $k = \frac{n}{2}$ the lemma is trivial. Suppose therefore that $k \geqq \frac{n+1}{2}$ and let $\xi = (k-\frac{\tau}{2})(n-k+1-\frac{\tau}{2})$. We note that condition (5.2.5) ensures that $\xi > 0$, $(k+1-\frac{\tau}{2})(n-k-\frac{\tau}{2}) > 0$ and $k(n-k+1-\tau) > 0$. Developing the left-hand side of the inequality in a binomial series we find

$$\left(\frac{(k-\frac{\tau}{2})(n-k+1-\frac{\tau}{2})}{(k+1-\frac{\tau}{2})(n-k-\frac{\tau}{2})}\right)^{\tau-1} = \left(1-\frac{2k-n}{\xi}\right)^{-(\tau-1)} = \sum_{i=0}^{\infty} c_i \left(\frac{2k-n}{\xi}\right)^i ,$$

where $c_i = \frac{1}{i!}(\tau-1)\tau(\tau+1)\ldots(\tau+i-2)$; the series converges as $0 < \frac{2k-n}{\xi} < 1$. Also

$$1 + (\tau-1)\frac{2k-n}{k(n-k+1-\tau)} = 1 + \frac{(\tau-1)(2k-n)}{\xi}\left(1-\frac{\tau}{2\xi}(2k-n+\frac{\tau-2}{2})\right)^{-1} =$$

$$= 1 + (\tau-1)\frac{2k-n}{\xi} + \sum_{i=2}^{\infty} c_i' \frac{(2k-n)\left(2k-n+\frac{\tau-2}{2}\right)^{i-1}}{\xi^i} ,$$

where $c_i' = (\tau-1)\left(\frac{\tau}{2}\right)^{i-1}$; the series converges since $0 < \frac{2k-n+\frac{\tau-2}{2}}{\xi} < 1$.

Now the first two terms of both series are equal, $c_2 = c_2'$, and for $i \geqq 2$

$$\frac{c_{i+1}}{c_i} - \frac{c_{i+1}'}{c_i'} = \frac{\tau+i-1}{i+1} - \frac{\tau}{2} = \frac{(2-\tau)(i-1)}{2(i+1)} .$$

Hence, for $1 < \tau \leqq 2$, $c_i' \leqq c_i$ and $0 < 2k-n+\frac{\tau-2}{2} \leqq 2k-n$,
whereas for $\tau \geqq 2$, $c_i' \geqq c_i$ and $0 < 2k-n < 2k-n+\frac{\tau-2}{2}$,
which proves the lemma.

Using this lemma we shall derive bounds for an incomplete beta integral. For all real τ and integer k and n satisfying (5.2.5) we define

$$(5.2.11) \qquad Q_{k,n}(\tau) = \int_{\frac{k-\frac{\tau}{2}}{n+1-\tau}}^{\frac{k+1-\frac{\tau}{2}}{n+1-\tau}} u^{-\tau}(1-u)^{-\tau}\,du \quad ,$$

and prove

LEMMA 5.2.8

For integer k and n satisfying (5.2.5)

$$Q_{k,n}(\tau) \leqq \frac{(n+1-\tau)^{2\tau-1}}{\left(k+\frac{1}{2}-\frac{\tau}{2}\right)^{\tau}\left(n-k+\frac{1}{2}-\frac{\tau}{2}\right)^{\tau}} \qquad \text{if } -1 \leqq \tau \leqq 0,$$

$$Q_{k,n}(\tau) \geqq \frac{(n+1-\tau)^{2\tau-1}}{\left(k+\frac{1}{2}-\frac{\tau}{2}\right)^{\tau}\left(n-k+\frac{1}{2}-\frac{\tau}{2}\right)^{\tau}} \qquad \text{if } 0 \leqq \tau \leqq 1,$$

$$Q_{k,n}(\tau) \geqq \frac{n(n+2-2\tau)}{k(n-k+1-\tau)} \frac{(n+1-\tau)^{2\tau-3}}{\left(k-\frac{\tau}{2}\right)^{\tau-1}\left(n-k+1-\frac{\tau}{2}\right)^{\tau-1}} \qquad \text{if } 1 < \tau \leqq \frac{3}{2}, \text{ and } k \geqq \frac{n}{2},$$

and

$$Q_{k,n}(\tau) \leqq \frac{n(n+2-2\tau)}{k(n-k+1-\tau)} \frac{(n+1-\tau)^{2\tau-3}}{\left(k-\frac{\tau}{2}\right)^{\tau-1}\left(n-k+1-\frac{\tau}{2}\right)^{\tau-1}} \qquad \text{if } \tau \geqq 2, \text{ and } k \geqq \frac{n}{2}.$$

PROOF

For $\tau \geqq 0$ the integrand $f(u) = u^{-\tau}(1-u)^{-\tau}$ of (5.2.11) is convex for $0 < u < 1$. Denoting by L its line of support at $u = \frac{k+\frac{1}{2}-\frac{\tau}{2}}{n+1-\tau}$ we obtain

$$(5.2.12)\quad Q_{k,n}(\tau) \geqq \int_{\frac{k-\frac{\tau}{2}}{n+1-\tau}}^{\frac{k+1-\frac{\tau}{2}}{n+1-\tau}} L(u)\,du = \frac{1}{n+1-\tau}\, L\left(\frac{k+\frac{1}{2}-\frac{\tau}{2}}{n+1-\tau}\right) =$$

$$= \frac{1}{n+1-\tau}\, f\left(\frac{k+\frac{1}{2}-\frac{\tau}{2}}{n+1-\tau}\right) = \frac{(n+1-\tau)^{2\tau-1}}{\left(k+\frac{1}{2}-\frac{\tau}{2}\right)^{\tau}\left(n-k+\frac{1}{2}-\frac{\tau}{2}\right)^{\tau}}\,, \quad \text{for } \tau \geqq 0.$$

For $-1 \leqq \tau \leqq 0$, f is concave on $(0,1)$ and inequality (5.2.12) is reversed.

The case where $\tau > 1$ and $k \geqq \frac{n}{2}$ remains to be considered. After multiplication of the integrand by $u^2 + 2u(1-u) + (1-u)^2 = 1$ we find by partial integration

$$(5.2.13)\quad \int u^{-\tau}(1-u)^{-\tau}\,du =$$

$$= \frac{1}{\tau-1}(2u-1)u^{-(\tau-1)}(1-u)^{-(\tau-1)} + \frac{4\tau-6}{\tau-1}\int u^{-(\tau-1)}(1-u)^{-(\tau-1)}\,du\,,$$

and since $u^{-(\tau-1)}(1-u)^{-(\tau-1)}$ is convex for $0 < u < 1$ we have

$$(5.2.14)\quad \int_{\frac{k-\frac{\tau}{2}}{n+1-\tau}}^{\frac{k+1-\frac{\tau}{2}}{n+1-\tau}} u^{-(\tau-1)}(1-u)^{-(\tau-1)}\,du \leqq$$

$$\leqq \frac{1}{2}\,\frac{(n+1-\tau)^{2\tau-3}}{\left(k-\frac{\tau}{2}\right)^{\tau-1}\left(n-k+1-\frac{\tau}{2}\right)^{\tau-1}} + \frac{1}{2}\,\frac{(n+1-\tau)^{2\tau-3}}{\left(k+1-\frac{\tau}{2}\right)^{\tau-1}\left(n-k-\frac{\tau}{2}\right)^{\tau-1}}\,.$$

Hence for $1 < \tau \leqq \frac{3}{2}$, $\frac{4\tau-6}{\tau-1} \leqq 0$ and by (5.2.13) and (5.2.14)

$$(5.2.15) \qquad Q_{k,n}(\tau) \geqq$$

$$\geqq -\frac{1}{\tau-1}\frac{(n+1-\tau)^{2\tau-3}(2k-n-1)}{\left(k-\frac{\tau}{2}\right)^{\tau-1}\left(n-k+1-\frac{\tau}{2}\right)^{\tau-1}}+\frac{1}{\tau-1}\frac{(n+1-\tau)^{2\tau-3}(2k-n+1)}{\left(k+1-\frac{\tau}{2}\right)^{\tau-1}\left(n-k-\frac{\tau}{2}\right)^{\tau-1}}+$$

$$+\frac{2\tau-3}{\tau-1}\frac{(n+1-\tau)^{2\tau-3}}{\left(k-\frac{\tau}{2}\right)^{\tau-1}\left(n-k+1-\frac{\tau}{2}\right)^{\tau-1}}+\frac{2\tau-3}{\tau-1}\frac{(n+1-\tau)^{2\tau-3}}{\left(k+1-\frac{\tau}{2}\right)^{\tau-1}\left(n-k-\frac{\tau}{2}\right)^{\tau-1}}=$$

$$=-\frac{(n+1-\tau)^{2\tau-3}(2k-n+2-2\tau)}{(\tau-1)\left(k-\frac{\tau}{2}\right)^{\tau-1}\left(n-k+1-\frac{\tau}{2}\right)^{\tau-1}}+\frac{(n+1-\tau)^{2\tau-3}(2k-n+2\tau-2)}{(\tau-1)\left(k+1-\frac{\tau}{2}\right)^{\tau-1}\left(n-k-\frac{\tau}{2}\right)^{\tau-1}}$$

As $\tau > 1$ and $k \geqq \frac{n}{2}$, $\frac{2k-n+2\tau-2}{\tau-1} > 0$ and as a result application of lemma 5.2.7 to the second term of the last member of (5.2.15) gives

$$(5.2.16) \quad Q_{k,n}(\tau) \geqq \frac{(n+1-\tau)^{2\tau-3}}{(\tau-1)\left(k-\frac{\tau}{2}\right)^{\tau-1}\left(n-k+1-\frac{\tau}{2}\right)^{\tau-1}}\ .$$

$$\cdot\left[-(2k-n+2-2\tau)+(2k-n+2\tau-2)\left\{1+(\tau-1)\frac{2k-n}{k(n-k+1-\tau)}\right\}\right]=$$

$$=\frac{(n+1-\tau)^{2\tau-3}}{\left(k-\frac{\tau}{2}\right)^{\tau-1}\left(n-k+1-\frac{\tau}{2}\right)^{\tau-1}}\left[4+\frac{(2k-n+2\tau-2)(2k-n)}{k(n-k+1-\tau)}\right]=$$

$$=\frac{n(n+2-2\tau)}{k(n-k+1-\tau)}\frac{(n+1-\tau)^{2\tau-3}}{\left(k-\frac{\tau}{2}\right)^{\tau-1}\left(n-k+1-\frac{\tau}{2}\right)^{\tau-1}} \quad \text{for } 1<\tau\leqq\frac{3}{2} \text{ and } k\geqq\frac{n}{2}\ .$$

For $\tau \geq 2$ the inequalities in (5.2.15) and lemma 5.2.7 are reversed and hence (5.2.16) is reversed too, which proves the lemma.

Lemma 5.2.9 summarizes the results that will be needed in the next section:

LEMMA 5.2.9

For integer k and n satisfying (5.2.5)

$$P_{k,n}(\tau) \leq Q_{k,n}(\tau) \qquad \text{if } 0 \leq \tau \leq \tfrac{3}{2}\,,$$

$$P_{k,n}(\tau) \geq Q_{k,n}(\tau) \qquad \text{if } -1 \leq \tau \leq 0\,,\ \text{or } \tau \geq 2\,.$$

PROOF

As $P_{k,n}(\tau) = P_{n-k,n}(\tau)$ and $Q_{k,n}(\tau) = Q_{n-k,n}(\tau)$ it is sufficient to prove the lemma for $k \geq \frac{n}{2}$. This is done by applying lemmata 5.2.6 and 5.2.8.

5.3. SMALL SAMPLE INEQUALITIES

Returning to the symmetric inverse beta distributions F_τ^* discussed in section 5.1 we may now prove

THEOREM 5.3.1

If $i \geq \frac{n+1}{2}$ and $E\,\underline{x}_{i:n}^*(\tau)$ exists, then

$$F_\tau^*\left(E\,\underline{x}_{i:n}^*(\tau)\right) \leq \frac{i-\frac{\tau}{2}}{n+1-\tau} \qquad \text{if } 0 \leq \tau \leq \tfrac{3}{2}\,,$$

$$F_\tau^*\left(E\,\underline{x}_{i:n}^*(\tau)\right) \geq \frac{i-\frac{\tau}{2}}{n+1-\tau} \qquad \text{if } -1 \leq \tau \leq 0\,,\ \text{or } \tau \geq 2\,.$$

PROOF

From section 5.1 we recall that the existence of $E\,\underline{x}^{*}_{i:n}(\tau)$ for $i \geq \frac{n+1}{2}$ implies that $i < n+2-\tau$ and that $E\,\underline{x}^{*}_{k:n}(\tau)$ exists for $\frac{n}{2} \leq k \leq i$. Inserting the abbreviation (5.2.6) in (5.1.7) we have for $\frac{n}{2} \leq k \leq i-1$

$$E\,\underline{x}^{*}_{k+1:n}(\tau) = E\,\underline{x}^{*}_{k:n}(\tau) + P_{k,n}(\tau) ,$$

and hence (cf. (5.1.8))

$$(5.3.1) \qquad E\,\underline{x}^{*}_{i:n}(\tau) = \sum_{k=\frac{n+1}{2}}^{i-1} P_{k,n}(\tau) \qquad \text{if n is odd, and}$$

$$(5.3.1') \qquad E\,\underline{x}^{*}_{i:n}(\tau) = \sum_{k=\frac{n}{2}+1}^{i-1} P_{k,n}(\tau) + \frac{1}{2}P_{\frac{n}{2},n}(\tau) \qquad \text{if n is even.}$$

Also by (5.1.5) and (5.2.11) for $\frac{n}{2} \leq k \leq i-1$

$$G^{*}_{\tau}\left(\frac{k+1-\frac{\tau}{2}}{n+1-\tau}\right) = G^{*}_{\tau}\left(\frac{k-\frac{\tau}{2}}{n+1-\tau}\right) + Q_{k,n}(\tau) ;$$

and since $G^{*'}_{\tau}$ is symmetric about $y = \frac{1}{2}$ and $G^{*}_{\tau}(\frac{1}{2}) = 0$,

$$(5.3.2) \qquad G^{*}_{\tau}\left(\frac{i-\frac{\tau}{2}}{n+1-\tau}\right) = \sum_{k=\frac{n+1}{2}}^{i-1} Q_{k,n}(\tau) \qquad \text{if n is odd, and}$$

$$(5.3.2') \qquad G^{*}_{\tau}\left(\frac{i-\frac{\tau}{2}}{n+1-\tau}\right) = \sum_{k=\frac{n}{2}+1}^{i-1} Q_{k,n}(\tau) + \frac{1}{2}Q_{\frac{n}{2},n}(\tau) \qquad \text{if n is even.}$$

Furthermore, if $\frac{n}{2} \leq k \leq i-1$, it follows from $\frac{n+1}{2} \leq i < n+2-\tau$ that

$\tau-1 < \frac{n}{2} \leqq k < n+1-\tau$ and hence the summation index k in (5.3.1), (5.3.1'), (5.3.2) and (5.3.2') satisfies condition (5.2.5). Application of lemma 5.2.9 completes the proof.

It may be useful to comment upon the limitations of theorem 5.3.1. In the first place at most one bound is provided for $F_\tau^*\left(E\ \underline{x}_{i:n}^*(\tau)\right)$ for any $\tau \neq 0$. It would seem, however, to demand an entirely different approach to establish a second bound, because this bound does not appear to be reached asymptotically even for $\frac{i}{n} \to 0$ or 1, but only for small samples to which the technique we have used is not very well adapted. In the second place the theorem does not give any results at all for $\tau < -1$ and $\frac{3}{2} < \tau < 2$. For $\tau < -1$ one might conjecture that the inequality for $-1 \leqq \tau \leqq 0$ will continue to hold, but these U-shaped distributions seem to be of only limited interest in practice. The distributions F_τ^* having $\frac{3}{2} < \tau < 2$ form a far more interesting part of the s-chain: these are distributions on $(-\infty, +\infty)$ having fairly heavy tails but still possessing a finite expectation. However, although it may be possible to prove the validity of the first inequality for values of τ somewhat greater than $\frac{3}{2}$, the difficulty here is that for values of τ near 2 neither inequality can hold for all $i \geqq \frac{n+1}{2}$. This may be seen as follows.

In [4] G. BLOM investigated the quantities

$$\alpha_i = \lim_{n\to\infty} \alpha_{i,n} = \lim_{n\to\infty} \alpha_{n+1-i:n}$$

for fixed i. Denoting these quantities by $\alpha_i^*(\tau)$ for the distribution F_τ^* it turns out that

$$\alpha_i^*(\tau) = i - \left(\frac{\Gamma(i+1-\tau)}{\Gamma(i)}\right)^{\frac{1}{1-\tau}} = 1 - a_{i-1}(2-\tau)$$

in our notation. By lemmata 5.2.2 and 5.2.4, $\alpha_i^*(\tau) < \frac{\tau}{2}$ for $0 < \tau < 2$. On the other hand $\alpha_{i:n}^*(2)$ is not identically equal to 1 for all

$i \geqq \frac{n+1}{2}$ and hence $\alpha^*_{i_0:n_0}(2) > 1$ for some $i_0 \geqq \frac{n_0+1}{2}$ by theorem 5.3.1 (e.g. $E\, \underline{x}^*_{3:4}(2) = 3 > 2.9 = G^*_2(\frac{2}{3})$). By continuity we have $\alpha^*_{i_0:n_0}(\tau) > \frac{\tau}{2}$ if τ is sufficiently near 2, $\tau < 2$, which shows that neither inequality of the theorem can hold.

Theorem 5.3.1 may now be applied to obtain small sample inequalities for distributions other than F^*_τ. Let $F \in \mathcal{S}$ be a given distribution function, let $\alpha_{i:n}$ and $\alpha(r)$ be defined by (5.1.1) and (5.1.2), and let

$$(5.3.3) \qquad \alpha_{inf} = \inf_{\frac{1}{2}<r<1} \alpha(r) = \inf_{\frac{1}{2}<r<1} \frac{r(1-r)}{2(2r-1)} \frac{G''(r)}{G'(r)} \text{ , and}$$

$$(5.3.4) \qquad \alpha_{sup} = \sup_{\frac{1}{2}<r<1} \alpha(r) = \sup_{\frac{1}{2}<r<1} \frac{r(1-r)}{2(2r-1)} \frac{G''(r)}{G'(r)} \text{ .}$$

Using the s-chain of symmetric inverse beta distributions F^*_τ as standards for s-comparison we find

THEOREM 5.3.2

If $F \in \mathcal{S}$, $i \geqq \frac{n+1}{2}$ and $E\, \underline{x}_{i:n}$ exists,

$$F(E\, \underline{x}_{i:n}) \leqq \frac{i-\alpha_{sup}}{n+1-2\alpha_{sup}} \qquad \text{if } 0 \leqq \alpha_{sup} \leqq \frac{3}{4} \text{ ,}$$

$$F(E\, \underline{x}_{i:n}) \geqq \frac{i-\alpha_{inf}}{n+1-2\alpha_{inf}} \qquad \text{if } -\frac{1}{2} \leqq \alpha_{inf} \leqq 0 \text{ , or } \alpha_{inf} \geqq 1 \text{ .}$$

PROOF

By lemma 4.4.2 and (5.1.6), $F \underset{s}{<} F^*_\tau$ if and only if $\alpha_{sup} \leqq \frac{\tau}{2}$, and $F^*_\tau \underset{s}{\leq} F$ if and only if $\alpha_{inf} \geq \frac{\tau}{2}$, hence

$$F^*_{2\alpha_{inf}} \underset{s}{\leq} F \underset{s}{\leq} F^*_{2\alpha_{sup}} .$$

(An equivalent way of formulating this is to remark that for F^*_τ , $\alpha^*(r) \equiv \frac{\tau}{2}$ for $\frac{1}{2} < r < 1$ and appeal to theorem 4.5.1). The theorem now follows from theorems 4.5.1 and 5.3.1.

We note that in theorem 5.3.2 we have only given those inequalities that are asymptotically sharp. The inequality $F(E\, \underline{x}_{i:n}) \leq \frac{i}{n+1}$ continues to hold if $\alpha_{sup} < 0$ and the reversed inequality holds for $\alpha_{inf} > 0$ too. This, however, is simply s-comparison with the uniform distribution F^*_0 which has already been discussed in paragraph 4.6.1.

The result of paragraph 4.6.3 concerning s-comparison of the normal distribution and the logistic distribution F^*_1 is a special case of the theorem since $\alpha_{sup} = \frac{1}{2}$ for the normal distribution. Other examples of this technique of generating small sample inequalities from their large sample counterparts are as follows.

5.3.1. CAUCHY'S DISTRIBUTION

For CAUCHY's distribution

$$F(x) = \frac{1}{2} + \frac{1}{\pi} \text{arctn } x \quad , \qquad -\infty < x < \infty ,$$

BLOM [4] gives $\alpha_{inf} = \alpha(1) = 1$ and $\alpha_{sup} = \alpha(\frac{1}{2}) = \frac{\pi^2}{8}$. For $i \neq 1$, $i \neq n$, $E\, \underline{x}_{i:n}$ exists, hence

$$F(E\, \underline{x}_{i:n}) \geq \frac{i-1}{n-1} \qquad \text{for } \frac{n+1}{2} \leq i \leq n-1 .$$

5.3.2. SYMMETRIC BETA DISTRIBUTIONS

The distribution functions

$$F(x) = \frac{\Gamma(2\tau)}{2^{2\tau-1}\left(\Gamma(\tau)\right)^2} \int_{-1}^{x} (1-t^2)^{\tau-1}\, dt \quad ,$$

$\tau > 0$, $-1 < x < 1$, are equivalent to the symmetric beta distributions (cf. paragraph 4.6.2). Substituting $r = F(x)$ in (5.1.2) we find

$$\alpha\left(F(x)\right) = \frac{F(x)\,(1-F(x))}{2(2F(x)-1)} \frac{G''F(x)}{G'F(x)} = -\frac{F(x)\,(1-F(x))}{2(2F(x)-1)} \frac{F''(x)}{F'^2(x)} =$$

$$= (\tau-1) \frac{x}{2F(x)-1} \frac{F(x)\,(1-F(x))}{(1-x^2)\,F'(x)} \quad ,$$

and by L'HÔPITAL's rule

$$\alpha(\tfrac{1}{2}) = \alpha\left(F(0)\right) = \frac{\tau-1}{4F'(0)} \lim_{x\to 0} \frac{x}{2F(x)-1} = \frac{\tau-1}{8F'^2(0)} = \frac{\tau-1}{2c(\tau)} \, ,$$

$$\alpha(1) = \alpha\left(F(1)\right) = \frac{\tau-1}{2} \lim_{x\to 1} \frac{1-F(x)}{(1-x)F'(x)} = \frac{\tau-1}{2\tau} \quad ,$$

where $$c(\tau) = 4F'^2(0) = \frac{\left(\Gamma(2\tau)\right)^2}{2^{4(\tau-1)}\left(\Gamma(\tau)\right)^4} = \frac{4\left(\Gamma(\tau+\frac{1}{2})\right)^2}{\pi\left(\Gamma(\tau)\right)^2}$$

by LEGENDRE's duplication formula. As $c(1) = 1$ and $\frac{c(\tau)}{\tau}$ is easily shown to be an increasing function of τ we have $\alpha(\frac{1}{2}) \leqq \alpha(1)$ for all $\tau > 0$. More generally one may prove by straightforward but somewhat lengthy algebra that $\alpha_{inf} = \alpha(\frac{1}{2})$ and $\alpha_{sup} = \alpha(1)$ for all $\tau > 0$ and hence for $i \geqq \frac{n+1}{2}$

$$F(E\,\underline{x}_{i:n}) \leqq \frac{i-\frac{\tau-1}{2\tau}}{n+1-\frac{\tau-1}{\tau}} \qquad \text{if} \quad \tau \geqq 1 ,$$

and

$$F(E\,\underline{x}_{i:n}) \geqq \frac{i-\frac{\tau-1}{2c(\tau)}}{n+1-\frac{\tau-1}{c(\tau)}} \qquad \text{if} \quad \tau \leqq 1,\ \tau + c(\tau) \geqq 1 .$$

Chapter 6

APPLICATIONS TO HYPOTHESIS TESTING AND ESTIMATION

6.1. COMPARISON OF NORMAL SCORES AND WILCOXON TESTS

It was pointed out in chapter 1 that the weak-order relations $\leq_c$ and $\leq_s$ may well be much better suited to express increasing skewness to the right and heavier tails than the standard measures of skewness and kurtosis. In this chapter we shall illustrate this point by considering some examples of comparison of distributions where these properties play an important part.

The first example - as well as the title of this section - is taken from a paper by J.L. HODGES jr. and E.L. LEHMANN [15]. They discuss the relative asymptotic efficiency $e_{W,N}(F)$ of WILCOXON's two sample test W to the normal scores test N, for the case where the underlying distribution is given by F. The following values of $e_{W,N}(F)$ are given by them.

TABLE 6.1.1

F	$e_{W,N}(F)$
Rectangular	0
Exponential	0
Normal	$3/\pi \approx 0.955$
Logistic	$\pi/3 \approx 1.05$
Double exponential	$3\pi/8 \approx 1.18$
Cauchy	1.413

They remark: "Qualitatively we may venture to guess that the Normal scores test is preferable when the distribution has an abrupt tail, like the rectangular; that they are about equally good with a bell-shaped density with a thin tail; and that the Wilcoxon test will perform relatively better when the tails are heavy so that the information is mainly to be found in the central rankings".

HODGES and LEHMANN also give an explicit formula for $e_{W,N}(F)$. Under certain regularity conditions they find

$$(6.1.1) \qquad e_{W,N}(F) = 12\left(\frac{\int_I F'^2(x)\,dx}{\int_I \frac{F'^2(x)}{F^{**'}G^{**}F(x)}\,dx} \right)^2$$

where F^{**} denotes the standard normal distribution function and G^{**} its inverse. By substitution of $x = G(y)$ we may rewrite (6.1.1) in a more convenient form

$$(6.1.2) \qquad e_{W,N}(F) = 12\left(\frac{\int_0^1 \frac{1}{G'(y)}\,dy}{\int_0^1 \frac{G^{**'}(y)}{G'(y)}\,dy} \right)^2 .$$

The regularity conditions given by HODGES and LEHMANN for (6.1.1) to hold are rather restrictive. Although it may be difficult to replace these conditions by less restrictive ones of any simplicity, it is clear that (6.1.1) and (6.1.2) will hold for a larger class of distributions than is indicated in [15]. We shall not, however, discuss this point here and we shall simply require (6.1.2) to hold for all distributions under consideration.

For symmetric distributions we prove a theorem that shows that heavier tails do indeed lead to larger values of $e_{W,N}(F)$.

THEOREM 6.1.1

If for F, $F^* \in \mathcal{S}$ expression (6.1.2) for $e_{W,N}$ is valid, then $F \underset{s}{<} F^*$ implies $e_{W,N}(F) \leqq e_{W,N}(F^*)$.

PROOF

We begin by remarking that, since HODGES and LEHMANN have shown that $e_{W,N}(F) \leqq \frac{6}{\pi}$, validity of formula (6.1.2) for $e_{W,N}(F)$ and $e_{W,N}(F^*)$ implies

$$\int_0^1 \frac{1}{G'(y)}\, dy < \infty \quad \text{and} \quad \int_0^1 \frac{1}{G^{*'}(y)}\, dy < \infty .$$

(Incidentally we note that $F \underset{s}{<} F^*$ implies that the latter integral is finite if the former is). Since both integrals are also positive we may suppose without loss of generality that F and F^* are standardized in such a way that

$$F(0) = F^*(0) = F^{**}(0) = \frac{1}{2} \tag{6.1.3}$$

and

$$\int_0^1 \frac{1}{G'(y)}\, dy = \int_0^1 \frac{1}{G^{*'}(y)}\, dy > 0 . \tag{6.1.4}$$

Now $F \underset{s}{<} F^*$ and hence by lemma 4.4.2, $\frac{G^{*'}(y)}{G'(y)}$ is non-decreasing for $\frac{1}{2} < y < 1$. Since it is also symmetrical about $y = \frac{1}{2}$ we find from (6.1.4)

$$\frac{1}{G'(y)} - \frac{1}{G^{*'}(y)} \leqq 0 \qquad \text{for } |y - \tfrac{1}{2}| \leqq c$$

and

$$\frac{1}{G'(y)} - \frac{1}{G^{*'}(y)} \geqq 0 \qquad \text{for } |y - \tfrac{1}{2}| \geqq c$$

for some $0 < c < \frac{1}{2}$. As F^{**} denotes a unimodal distribution, $G^{**'}$ is non-decreasing for $\frac{1}{2} < y < 1$, and as it is also symmetrical about $y = \frac{1}{2}$ we have

$$G^{**'}(y) \leq G^{**'}(\tfrac{1}{2}+c) \qquad \text{for } |y-\tfrac{1}{2}| \leq c$$

and

$$G^{**'}(y) \geq G^{**'}(\tfrac{1}{2}+c) \qquad \text{for } |y-\tfrac{1}{2}| \geq c \quad .$$

Hence

$$\int_0^1 G^{**'}(y)\left(\frac{1}{G'(y)} - \frac{1}{G^{*'}(y)}\right)dy \geq G^{**'}(\tfrac{1}{2}+c)\int_0^1\left(\frac{1}{G'(y)} - \frac{1}{G^{*'}(y)}\right)dy = 0,$$

or

$$(6.1.5) \qquad 0 < \int_0^1 \frac{G^{**'}(y)}{G^{*'}(y)}\,dy \leq \int_0^1 \frac{G^{**'}(y)}{G'(y)}\,dy \quad ,$$

where the latter integral or both integrals may diverge to $+\infty$. Combination of (6.1.2), (6.1.4) and (6.1.5) completes the proof.

It may be appropriate to remark that a theorem like 6.1.1 is, of course, not an isolated result. Bearing in mind that WILCOXON's test is the locally most powerful rank test for the logistic distribution we see that the comparison really involves four distributions: F, F^{*}, the logistic and the normal distributions. A tentative study shows that analogous results may be obtained for distributions other than logistic and normal, and further research in this direction will be undertaken.

6.2. STUDENT'S TEST UNDER NON-STANDARD CONDITIONS

The second example in this chapter is taken from a paper by H. HOTELLING [17]. Let $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_n$ be independent and identically distributed with common distribution function $F \in \mathcal{F}$, for which either $\mu = E\,\underline{x}$ exists, or $F \in \mathcal{S}$; in the latter case we define μ by $F(\mu) = \frac{1}{2}$. Furthermore let

$$\underline{t}_n = \frac{\overline{\underline{x}} - \mu}{\underline{s}} \sqrt{n} ,$$

where $\overline{\underline{x}} = \frac{1}{n} \sum_{i=1}^{n} \underline{x}_i$ and $\underline{s}^2 = \frac{1}{n-1} \sum_{i=1}^{n} (\underline{x}_i - \overline{\underline{x}})^2$.

The probability that $\underline{t}_n$ will exceed a constant value t will be denoted by $P(\underline{t}_n > t \mid F)$ and we define

$$(6.2.1) \qquad R_n(F) = \lim_{t \to \infty} \frac{P(\underline{t}_n > t \mid F)}{P(\underline{t}_n > t \mid F^{**})} ,$$

where F^{**} denotes the normal distribution function.

Suppose that, assuming the underlying distribution to be normal, one carries out STUDENT's right-sided test for the hypothesis $\mu \leqq \mu_0$, whereas in fact F is not normal at all. Then obviously $R_n(F)$ denotes the limit of the ratio of the actual size to the assumed size of the test as both these sizes tend to zero. It may therefore serve to provide a rough idea of what to expect when the assumption of normality is violated.

In [17] HOTELLING showed that

$$(6.2.2) \qquad R_n(F) = \frac{2(\pi n)^{\frac{n}{2}}}{\Gamma\left(\frac{n}{2}\right)} \int_{\mu}^{\infty} [F'(x)]^n \, (x-\mu)^{n-1} \, dx ,$$

which we may rewrite in terms of G as

$$(6.2.3) \qquad R_n(F) = \frac{2(\pi n)^{\frac{n}{2}}}{\Gamma(\frac{n}{2})} \int_{F(\mu)}^{1} \left[\frac{G(y)-\mu}{G'(y)}\right]^{n-1} dy \quad .$$

For $n = 3$ HOTELLING found $R_3(F) = 1$, 0.785 and 0.413 for the normal, double exponential and CAUCHY distributions respectively, which seems to indicate - paradoxally enough at first sight - the $R_n(F)$ decreases as the tails of F become heavier. Theorem 6.2.1 shows this idea to be correct for s-ordered symmetric distributions; moreover, the same result is proved for c-ordered distributions.

THEOREM 6.2.1

If $F, F^* \in \mathcal{F}$, and if either $E\,\underline{x}$, $E\,\underline{x}^*$ exist and $F \underset{c}{<} F^*$, or $F, F^* \in \mathcal{S}$ and $F \underset{s}{<} F^*$, then

$$R_n(F) \geqq R_n(F^*) \qquad \text{for } n = 2,3,\ldots \quad .$$

PROOF

Without loss of generality we may set $\mu = \mu^* = 0$, where μ and μ^* are either expectation or point of symmetry of F and F^* respectively.

Suppose first that $F, F^* \in \mathcal{S}$ and hence that $G(\frac{1}{2}) = G^*(\frac{1}{2}) = 0$. From $F \underset{s}{<} F^*$ we find by lemma 4.4.2 that

$$\frac{G^{*'}(y)}{G'(y)} \text{ is non-decreasing for } \frac{1}{2} < y < 1$$

and as a result (we note that G and G^* are increasing)

$$\frac{G^*(y)}{G(y)} \leqq \frac{G^{*'}(y)}{G'(y)} \qquad \text{for } \frac{1}{2} < y < 1 \ , \qquad \text{or}$$

$$(6.2.4) \qquad 0 \leqq \frac{G^*(y)}{G^{*'}(y)} \leqq \frac{G(y)}{G'(y)} \quad \text{for } \frac{1}{2} < y < 1 \quad .$$

Application of (6.2.4) to (6.2.3) for $\mu = \mu^* = 0$, $F(\mu) = F^*($ proves the theorem for the case of s-ordering.

If $F \underset{c}{\leq} F^*$ we have

$$F(0) = F(E\,\underline{x}) \leqq F^*(E\,\underline{x}^*) = F^*(0)$$

by theorem 4.2.1 and by lemma 4.1.3

$$\frac{G^{*'}(y)}{G'(y)} \text{ is non-decreasing for } 0 < y < 1 \quad .$$

Hence for $F(0) \leqq F^*(0) < y < 1$, $G(y)$ and $G^*(y)$ are positive

$$\frac{G^*(y)}{G(y)} \leqq \frac{G^*(y)}{G(y) - GF^*(0)} = \frac{G^*(y) - G^*F^*(0)}{G(y) - GF^*(0)} \leqq \frac{G^{*'}(y)}{G'(y)}$$

$$(6.2.5) \qquad 0 \leqq \frac{G^*(y)}{G^{*'}(y)} \leqq \frac{G(y)}{G'(y)} \quad \text{for } F^*(0) < y < 1$$

Consequently

$$\int_{F^*(0)}^{1} \left[\frac{G^*(y)}{G^{*'}(y)}\right]^{n-1} dy \leqq \int_{F^*(0)}^{1} \left[\frac{G(y)}{G'(y)}\right]^{n-1} dy \leqq \int_{F(0)}^{1} \left[\frac{G(y}{G'(}\right.$$

since the latter integrand is positive for $y > F(0)$. This com the proof.

6.3. EFFICIENCY OF MEDIAN AND MEAN

Having discussed two applications to hypothesis testing in the previous sections we end by proving a result in estimation theory. Let $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_n$ denote a random sample from a symmetric distribution $F \in \mathcal{S}$ with finite variance $\sigma^2(\underline{x})$, and suppose one wishes to estimate $E\,\underline{x}$. Two unbiased estimators that are generally used in this situation are the sample median

$$\underline{x}_{\frac{n+1}{2}:n}$$

and the sample mean

$$\bar{\underline{x}}_n = \frac{1}{n}\sum_{i=1}^{n} \underline{x}_i \quad ,$$

where we have supposed n to be odd. The choice between them should depend on the ratio of their (small sample) efficiencies

$$(6.3.1) \qquad r_n(F) = \frac{\mathrm{eff}\left(\underline{x}_{\frac{n+1}{2}:n}\right)}{\mathrm{eff}\left(\bar{\underline{x}}_n\right)} = \frac{\sigma^2(\underline{x})}{n.\sigma^2\left(\underline{x}_{\frac{n+1}{2}:n}\right)} \quad .$$

Theorem 6.3.1 shows that $r_n(F)$ increases as the tails of the distribution become heavier, and hence that the median is only to be preferred as an estimator for distributions with high kurtosis. This result supports the statement by G.W. BROWN and J.W. TUKEY [5] that "it is probable that the relative efficiencies of mean and median are greatly affected by the length of the tail".

THEOREM 6.3.1

For distributions $F, F^* \in \mathcal{S}$ having finite variances $\sigma^2(\underline{x})$ and $\sigma^2(\underline{x}^*)$, $F \underset{s}{<} F^*$ implies

$$r_n(F) \leqq r_n(F^*) \qquad \text{for } n = 1,3,5,\dots \ .$$

PROOF

Without loss of generality we set $E\,\underline{x} = E\,\underline{x}^* = 0$, $\sigma^2(\underline{x}) = \sigma^2(\underline{x}^*)$. Since $\underline{x}^*$ is isomorous with $\phi(\underline{x}) = G^*F(\underline{x})$ and G^*F is antisymmetrical, concave-convex on I about $E\,\underline{x} = 0$ we know from the proof of theorem 2.3.2 that

$$\phi(x) = G^*F(x) \leqq x \qquad \text{for } 0 \leqq x \leqq x_0' \qquad \text{and}$$
$$\phi(x) = G^*F(x) \geqq x \qquad \text{for } x \geqq x_0' \ ,\ x \in I \ ,$$

for some $x_0' \geqq 0$, $x_0' \in I$, or

$$\text{(6.3.2)} \qquad \begin{aligned} 0 \leqq G^*(y) \leqq G(y) \qquad & \text{for } \tfrac{1}{2} \leqq y \leqq y_0 = F(x_0'),\ \text{and} \\ G^*(y) \geqq G(y) \geqq 0 \quad & \text{for } y_0 \leqq y < 1 \ . \end{aligned}$$

By (3.1.4) we have

$$\sigma^2\left(\underline{x}_{\frac{n+1}{2}:n}\right) - \sigma^2\left(\underline{x}^*_{\frac{n+1}{2}:n}\right) = \int_0^1 \left[G^2(y) - G^{*2}(y)\right] b_{\frac{n+1}{2}:n}(y)\,dy =$$

$$= 2\int_{\frac{1}{2}}^1 \left[G^2(y) - G^{*2}(y)\right] b_{\frac{n+1}{2}:n}(y)\,dy \geqq$$

$$\geqq 2\, b_{\frac{n+1}{2}:n}(y_0) \int_{\frac{1}{2}}^1 \left[G^2(y) - G^{*2}(y)\right] dy =$$

$$= b_{\frac{n+1}{2}:n}(y_0)\left[\sigma^2(\underline{x}) - \sigma^2(\underline{x}^*)\right] = 0 \ ,$$

since $b_{\frac{n+1}{2}:n}$ is symmetric about $y = \frac{1}{2}$ and non-increasing for $y \geqq \frac{1}{2}$.

Hence

$$\sigma^2\left(x^*_{\frac{n+1}{2}:n}\right) \leqq \sigma^2\left(x_{\frac{n+1}{2}:n}\right)$$

which proves the theorem.

REFERENCES

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

[2] A. BENARD, E.C. BOS-LEVENBACH — Het uitzetten van waarnemingen op waarschijnlijkheidspapier, Statistica Neerl. 7 (1953), 163-173.

[3] A.R. BLOEMENA — Sampling from a graph, Mathematical Centre Tracts 2, Mathematisch Centrum Amsterdam, 1964.

[4] G. BLOM — Statistical estimates and transformed beta-variables, John Wiley & Sons New York; Almqvist and Wiksell Stockholm, 1958.

[5] G.W. BROWN, J.W. TUKEY — Some distributions of sample means, Ann. Math. Statist. 17 (1946), 1-12.

[6] R. GARVER — Concerning the limits of a measure of skewness, Ann. Math. Statist. 3 (1932), 358-360.

[7] H. CHERNOFF, G.J. LIEBERMAN — Use of normal probability paper, Journ. Am. Stat. Ass. 49 (1954), 778-785.

[8] H. CRAMÉR — Mathematical methods of statistics, Princeton University Press, 1946.

[9] D. VAN DANTZIG, J. HEMELRIJK — Statistical methods based on few assumptions, Bull. Inst. Int. Stat. 34 (1954), 239-267.

[10] F.N. DAVID, N.L. JOHNSON — Statistical treatment of censored data, I: Fundamental formulae, Biometrika 41 (1954), 228-240.

[11] S.S. GUPTA — Order statistics from the gamma distribution, Technometrics 2 (1960), 243-262.

[12] J.B.S. HALDANE — Mode and median of a nearly normal distribution with given cumulants, Biometrika 32 (1942), 294-299.

[13] G.H. HARDY, J.E. LITTLEWOOD, G. PÓLYA — Inequalities, Cambridge Univ. Press, 1934.

[14] H.L. HARTER — Expected values of normal order statistics, Biometrika 48 (1961), 151-165.

[15] J.L. HODGES JR., E.L. LEHMANN Comparison of the normal scores and Wilcoxon tests, Proc. 4th Berkeley Symp. (1961) Vol. I, 307-317.

[16] W. HOEFFDING On the distribution of the expected values of order statistics, Ann. Math. Statist. 24 (1953), 93-100.

[17] H. HOTELLING The behavior of some standard statistical tests under non-standard conditions, Proc. 4th Berkeley Symp. (1961) Vol. I, 319-359.

[18] H. HOTELLING, L.M. SOLOMONS The limits of a measure of skewness, Ann. Math. Statist. 3 (1932), 141-142.

[19] J.L.W.V. JENSEN Sur les fonctions convexes et les inégalités entre les valeurs moyennes, Acta Math. 30 (1906), 175-193.

[20] M.G. KENDALL, A. STUART The advanced theory of statistics Vol. I, Griffin London, 1958.

[21] M. LOÈVE Probability theory, 3rd Ed., van Nostrand New York, 1963.

[22] K. PEARSON — Contributions to the mathematical theory of evolution II: Skew variation in homogeneous material, Philos. Trans. 186 (1895), 343-414; Karl Pearson's early statistical papers, Cambridge Univ. Press, 1948.

[23] A.E. SARHAN, B.G. GREENBERG, et al. — Contributions to order statistics, Wiley New York, 1962.

[24] D. TEICHROEW — Tables of expected values of order statistics and products of order statistics for samples of size twenty and less from the normal distribution, Ann. Math. Statist. 27 (1956), 410-426.

[25] E.T. WHITTAKER, G.S. WATSON — A course of modern analysis, Cambridge Univ. Press, 2nd Ed., 1915.

[26] A.C. ZAANEN — An introduction to the theory of integration, North Holland Publishing Cy Amsterdam, 1958.

[27] W.R. VAN ZWET — Two weak-order relations for distribution functions, Report S 303 (VP 17) of the Dept. of Math. Stat. of the Mathematical Centre Amsterdam; Ann. Math. Statist. 33 (1962), 1484.