Z. Wahrscheinlichkeitstheorie verw. Gebiete
36, 285–294 (1976)

Zeitschrift für
Wahrscheinlichkeitstheorie
und verwandte Gebiete

Inequalities for $\mathscr{E}k(X, Y)$ when the Marginals are Fixed

Stamatis Cambanis[1]*, Gordon Simons[1]*, and William Stout[2]**
[1] Department of Statistics, University of North Carolina, Chapel Hill, North Carolina 27514, USA
[2] Department of Mathematics, University of Illinois, Champaign-Urbana, Illinois 61820, USA

When $k(x, y)$ is a quasi-monotone function and the random variables X and Y have fixed distributions, it is shown under some further mild conditions that $\mathscr{E}k(X, Y)$ is a monotone functional of the joint distribution function of X and Y. Its infimum and supremum are both attained and correspond to explicitly described joint distribution functions.

1. Introduction

It is well known that if X and X' are random variables, if X is stochastically smaller than X' and if $k(x)$ is a monotone function then $k(X)$ is stochastically smaller than $k(X')$ and

$$\mathscr{E}k(X) \leqq \mathscr{E}k(X'),$$

and the ordering is reversed when $k(x)$ is an antitone function (see for instance p. 159 of Hardy, Littlewood and Pólya [6] and p. 179 of Veinott [15]). We say that X is stochastically smaller than X', and write $X \subset X'$, if $\Pr\{X<x\} \geqq \Pr\{X'<x\}$ for all x. Also $k(x)$ is called monotone (resp. antitone) if $k(x) \leqq k(x')$ (resp. $k(x) \geqq k(x')$) for all $x \leqq x'$.

We are interested here in a two-dimensional analog of this result. For pairs of random variables (X, Y) and (X', Y') we say that (X, Y) is stochastically smaller than (X', Y'), and we write $(X, Y) \subset (X', Y')$, if

$$\Pr\{X<x, Y<y\} \geqq \Pr\{X'<x, Y'<y\}$$

for every x and y. It is easy to see that the condition $(X, Y) \subset (X', Y')$ alone does not imply

$$\mathscr{E}k(X, Y) \leqq \mathscr{E}k(X', Y') \qquad (1)$$

* Research supported by the Air Force Office of Scientific Research under Grant AFOSR-75-2796
** Research supported by the National Science Foundation

for monotone functions $k(x, y)$ of two variables. A simple counter example is given by Veinott [15]. He goes on to show that (1) holds for monotone functions if certain additional conditions, involving the conditional distributions of Y given $X=x$ and of Y' given $X'=x'$, are imposed. (In fact this result is n-dimensional, $n \geqq 2$. For a slightly weaker set of conditions see Pledger and Proschan [11].) Here we are not concerned with monotone functions $k(x, y)$ but with quasi-monotone functions, which are analogs to monotone functions of one variable as well. The class of quasi-monotone (and quasi-antitone) functions contains many important functions which are not monotone (nor antitone). Many examples are given in Section 4. Just as the condition $(X, Y) \subset (X', Y')$ is insufficient to guarantee (1) or its reverse for monotone functions $k(x, y)$, it is insufficient for quasi-monotone functions as well. The major additional requirement we impose is that the corresponding marginal distributions of (X, Y) and (X', Y') are the same. This requirement, while strong, is both natural and necessary.

2. The Main Result

Consider the random variables X and Y defined on a probability space $(\Omega, \mathscr{F}, P)$ and let $F(x)$ and $G(y)$ be their distribution functions respectively and $H(x, y)$ their joint distribution function. In studying the dependence of $\mathscr{E} k(X, Y)$ on H it would be useful to have an appropriate expression, other than its definition $\int_{R^2} k(x, y)\, dH(x, y)$, and this is now done when k is quasi-monotone (or quasi-antitone). In order to produce slightly simpler expressions we assume all distribution functions to be left continuous.

A function $k(x, y)$ is called quasi-monotone if for all $x \leqq x'$ and $y \leqq y'$

$$\Delta_{(x, x')}^{(y, y')} k \equiv k(x, y) + k(x', y') - k(x, y') - k(x', y) \geqq 0,$$

and quasi-antitone if $\Delta_{(x, x')}^{(y, y')} k \leqq 0$, i.e. if $-k$ is quasi-monotone. If k is quasi-monotone and right continuous then it determines uniquely a (σ-finite, nonnegative) measure μ on the Borel subsets $\mathscr{B}^2$ of the plane R^2 such that for all $x \leqq x'$ and $y \leqq y'$,

$$\mu\{(x, x'] \times (y, y']\} = \Delta_{(x, x')}^{(y, y')} k, \tag{2}$$

(see p. 167 of von Neumann [10]). An interchange of the order of integration in an appropriate double integral gives then the desired expression for $\mathscr{E} k(X, Y)$.

Let us first illustrate the method in a particular case and then proceed to the general case. Let $k(x, y)$ be symmetric, right continuous and quasi-monotone. Define the function $f(x, y, \omega)$ by

$$f(x, y, \omega) = \begin{cases} 1 & \text{if } X(\omega) < x,\ y \leqq Y(\omega) \text{ or } Y(\omega) < x,\ y \leqq X(\omega) \\ 0 & \text{otherwise}. \end{cases}$$

Then f is clearly a measurable function on $(R^2 \times \Omega, \mathscr{B}^2 \times \mathscr{F}, \mu \times P)$ and since it is nonnegative we have by Fubini's theorem

$$\mathscr{E} \int_{R^2} f\, d\mu = \int_{R^2} \mathscr{E} f\, d\mu.$$

But clearly

$$\begin{aligned}
&\textstyle\int_{R^2} f\,d\mu = k(X,X)+k(Y,Y)-2k(X,Y),\\
&\mathscr{E}f = P\{X<x\wedge y,\ Y\geqq x\vee y\}+P\{X\geqq x\vee y,\ Y<x\wedge y\}\\
&\quad = F(x\wedge y)+G(x\wedge y)-H(x\vee y, x\wedge y)-H(x\wedge y, x\vee y)\\
&\quad = A(x,y), \quad \text{say,}
\end{aligned}\tag{3}$$

and thus

$$\mathscr{E}\{k(X,X)+k(Y,Y)-2k(X,Y)\}=\textstyle\int_{R^2} A\,d\mu. \tag{4}$$

If the expected values $\mathscr{E}k(X,X)$ and $\mathscr{E}k(Y,Y)$ are finite we obtain the desired expression

$$2\mathscr{E}k(X,Y)=\mathscr{E}k(X,X)+\mathscr{E}k(Y,Y)-\textstyle\int_{R^2} A\,d\mu. \tag{5}$$

Now if the marginal distributions F and G are fixed, $\mathscr{E}k(X,Y)$ depends on H only through A, and since μ is a nonnegative measure, increasing H will result in increasing $\mathscr{E}k(X,Y)$; in other words $\mathscr{E}k(X,Y)$ is a monotone functional of H. (And it is clearly an antitone functional of H when k is an antitone function.)

The general case of a quasi-monotone, right continuous function $k(x,y)$ requires a different choice of f, but the idea is of course the same. The appropriate function $f(x,y,\omega)$ is now defined by

$$f(x,y,\omega)=\begin{cases}+1 & \text{if } x_0<x\leqq X(\omega),\ y_0<y\leqq Y(\omega)\\ & \text{or } X(\omega)<x\leqq x_0,\ Y(\omega)<y\leqq y_0\\ -1 & \text{if } X(\omega)<x\leqq x_0,\ y_0<y\leqq Y(\omega)\\ & \text{or } x_0<x\leqq X(\omega),\ Y(\omega)<y\leqq y_0\\ \ \ 0 & \text{otherwise}\end{cases}$$

where x_0 and y_0 are (appropriate) fixed real numbers. Again f is $\mathscr{B}^2\times\mathscr{F}$-measurable. If f^+ and f^- are the positive and negative parts of f, then by Fubini's theorem we have

$$\mathscr{E}\textstyle\int_{R^2} f^{\pm}\,d\mu=\int_{R^2}\mathscr{E}f^{\pm}\,d\mu.$$

Proceeding as before, and introducing

$$k_0(x,y)=k(x,y)-k(x,y_0)-k(x_0,y)+k(x_0,y_0), \tag{6}$$

we obtain

$$\mathscr{E}k_0^+(X,Y)=\textstyle\int_{R^2} B^+\,d\mu \quad\text{and}\quad \mathscr{E}k_0^-(X,Y)=\int_{R^2} B^-\,d\mu,$$

where k_0^+ and k_0^- are the positive and negative parts of k_0, and

$$\begin{aligned}
B^+(x,y)=\mathscr{E}f^+&=\begin{cases}1+H(x,y)-F(x)-G(y) & \text{if } x_0<x,\ y_0<y\\ H(x,y) & \text{if } x\leqq x_0,\ y\leqq y_0\\ 0 & \text{otherwise}\end{cases}\\
B^-(x,y)=\mathscr{E}f^-&=\begin{cases}F(x)-H(x,y) & \text{if } x\leqq x_0,\ y_0<y\\ G(y)-H(x,y) & \text{if } x_0<x,\ y\leqq y_0\\ 0 & \text{otherwise.}\end{cases}
\end{aligned}\tag{7}$$

If Z^+ and Z^- are the positive and negative parts of a random variable Z, using the standard terminology, we say that its expectation $\mathscr{E}Z$ exists (even if infinite valued) if at least one of $\mathscr{E}Z^+$ and $\mathscr{E}Z^-$ is finite, and then $\mathscr{E}Z$ is defined by $\mathscr{E}Z^+ - \mathscr{E}Z^-$. Thus, if the expectation $\mathscr{E}k_0(X, Y)$ exists (even if infinite valued) we have

$$\mathscr{E}k_0(X, Y) = \int_{R^2} d\mu \tag{8}$$

where $B = B^+ - B^-$. By (6), this is guaranteed if $\mathscr{E}k(X, Y)$ exists (even if infinite valued) and $\mathscr{E}k(X, y_0)$, $\mathscr{E}k(x_0, Y)$ are finite, in which case we have

$$\mathscr{E}k(X, Y) = \mathscr{E}k(X, y_0) + \mathscr{E}k(x_0, Y) - k(x_0, y_0) + \int_{R^2} B\, d\mu. \tag{9}$$

Hence for fixed F and G, $\mathscr{E}k(X, Y)$ depends on H only through B and thus it is a monotone functional of H (antitone when k is antitone).

Let us now summarize our results. We write $X \stackrel{d}{=} X'$ when the random variables X and X' have the same distributions.

Theorem 1. *Let* $X \stackrel{d}{=} X'$, $Y \stackrel{d}{=} Y'$ *and* $(X, Y) \subset (X', Y')$. *If* $k(x, y)$ *is a quasi-monotone, right continuous function then*

$$\mathscr{E}k(X, Y) \geqq \mathscr{E}k(X', Y') \tag{10}$$

when the expectations in (10) *exist (even if infinite valued) and either of the following is satisfied:*

(i) $k(x, y)$ *is symmetric and the expectations* $\mathscr{E}k(X, X)$ *and* $\mathscr{E}k(Y, Y)$ *are finite,*

(ii) *the expectations* $\mathscr{E}k(X, y_0)$ *and* $\mathscr{E}k(x_0, Y)$ *are finite for some* x_0 *and* y_0.

Of course the reverse inequality in (10) holds if k is quasi-antitone. It should be noted that the expression in the general case, given by (9) and (7), is greatly simplified when k is a multiple of a distribution function by taking the reference points x_0, y_0 at $-\infty$. However, in most interesting examples k is not a multiple of a distribution function (in fact μ is not even finite).

It should be remarked that $(X, Y) \subset (X', Y')$ and k quasi-monotone do not necessarily imply that $k(X, Y) \supset k(X', Y')$, from which (10) would follow trivially. This is seen by the following example. Let $(X, Y) \stackrel{d}{=} (B, B)$ and $(X', Y') \stackrel{d}{=} (B, B')$ where B, B' are independent Bernoulli random variables each with mean $\frac{1}{2}$, and let $k(x, y) = (x+y)^2$ (quasi-monotone). It is easily seen that $(X, Y) \subset (X', Y')$, $\mathscr{E}k(X, Y) = 2 > 1.5 = \mathscr{E}k(X', Y')$ and that neither $k(X, Y) \supset k(X', Y')$ nor $k(X, Y) \subset k(X', Y)$ is true.

The result as stated in Theorem 1 is not in its weakest form. It may be that inequality (10) holds assuming only that the expectations in (10) exist (even if infinite valued); however, at present, we can neither prove nor disprove this statement. Instead we offer the following remarks. When k is symmetric, introducing

$$k_s(x, y) = k(x, x) + k(y, y) - 2k(x, y) \geqq 0,$$

we have

$$\begin{aligned} 2k(x, y) &= k(x, x) + k(y, y) - k_s(x, y) \\ &= k^+(x, x) + k^+(y, y) - k^-(x, x) - k^-(y, y) - k_s(x, y) \end{aligned}$$

and from (3) and (4),

$$0 \leqq \mathscr{E}k_s(X, Y) \leqq \mathscr{E}k_s(X', Y') \leqq +\infty.$$

It then follows that assumption (i) may be weakened to

(i)′ *$k(x, y)$ is symmetric and either the expectations $\mathscr{E}k^+(X, X)$ and $\mathscr{E}k^+(Y, Y)$ are finite, or the expectations $\mathscr{E}k^-(X, X)$, $\mathscr{E}k^-(Y, Y)$, and $\mathscr{E}k_s(X, Y)$ are finite.*

Similarly condition (ii) for general k may be omitted whenever an appropriate truncation argument can be applied. At present we do not have a truncation argument to eliminate (ii) altogether. However the following truncation arguments clearly work for some specific k's mentioned in Section 4. It is clear that if all random variables X, Y, X', Y' are bounded and if k is locally bounded than (ii) is satisfied. Now for a function $f(x)$ and $c>0$, define $f^c(x) = -c$ if $f(x) < -c$, $= f(x)$ if $-c \leqq f(x) \leqq c$, $= c$ if $c < f(x)$. Notice that from (6) we have $k(x, y) = k_0(x, y) + f(x) + g(y)$ where $f(x) = k(x, y_0)$ and $g(y) = k(x_0, y) - k(x_0, y_0)$. For $c>0$ let

$$k^c(x, y) = k_0(x^c, y^c) + f^c(x) + g^c(y).$$

Then as $c \uparrow \infty$ we have $k^c(x, y) \to k(x, y)$ as well as $k(x^c, y^c) \to k(x, y)$ (with $k_0(x^c, y^c) \uparrow k_0(x, y)$ on the first and third quadrants of the plane around (x_0, y_0) and $k_0(x^c, y^c) \downarrow k_0(x, y)$ on the second and fourth quadrants). Assuming k is locally bounded, we clearly have

$$\mathscr{E}k^c(X, Y) \geqq \mathscr{E}k^c(X', Y') \quad \text{and} \quad \mathscr{E}k(X^c, Y^c) \geqq \mathscr{E}k(X'^c, Y'^c)$$

and thus assumption (ii) may be replaced by

(ii)′ *k is locally bounded and such that as $c \uparrow \infty$, either $\mathscr{E}k^c(X, Y) \to \mathscr{E}k(X, Y)$ and $\mathscr{E}k^c(X', Y') \to \mathscr{E}k(X', Y')$ for some x_0 and y_0, or $\mathscr{E}k(X^c, Y^c) \to \mathscr{E}k(X, Y)$ and $\mathscr{E}k(X'^c, Y'^c) \to \mathscr{E}k(X', Y')$.*

Condition (ii)′ is easily seen to be satisfied for several simple k's, e.g. xy, $(x+y)^2$, etc.

3. An Application to the Set of Values of $\mathscr{E}k(X, Y)$

In this section we denote by $\mathscr{H}(F, G)$ the class of all joint distribution functions $H(x, y)$ with fixed marginals $F(x)$ and $G(y)$, and by $\mathscr{E}_H k(X, Y)$ the expected value when H is the joint distribution function of X and Y. It is well known that $\mathscr{H}(F, G)$ has an upper and lower bound. In fact a distribution function $H(x, y)$ belongs to $\mathscr{H}(F, G)$ if and only if

$$H_-(x, y) \leqq H(x, y) \leqq H_+(x, y)$$

for all x and y, where the distribution functions H_- and H_+ are given by

$$H_-(x, y) = \max\{F(x) + G(y) - 1, 0\}, \qquad H_+(x, y) = \min\{F(x), G(y)\}$$

(see Hoeffding [8] and Fréchet [5]). $\mathscr{H}(F, G)$ is clearly a convex family of distribution functions.

Now let $\mathscr{H}$ be any convex family of bivariate distribution functions. If $H, H' \in \mathscr{H}$ are such that $\mathscr{E}_H k(X, Y)$ and $\mathscr{E}_{H'} k(X, Y)$ exist and are finite then each number in the closed (bounded) interval with endpoints $\mathscr{E}_H k(X, Y)$ and $\mathscr{E}_{H'} k(X, Y)$ is equal to $\mathscr{E}_{H''} k(X, Y)$ for some $H'' \in \mathscr{H}$. Indeed, if for each $\alpha \in [0, 1]$ we define $H_\alpha = \alpha H + (1-\alpha) H'$, then $H_\alpha \in \mathscr{H}$ and the conclusion follows from

$$\mathscr{E}_{H_\alpha} k(X, Y) = \alpha \mathscr{E}_H k(X, Y) + (1-\alpha) \mathscr{E}_{H'} k(X, Y).$$

This conclusion is no longer valid if

$$-\infty < \mathscr{E}_H k(X, Y) < \mathscr{E}_{H'} k(X, Y) = +\infty$$

as is seen by the example $\mathscr{H} = \{H_\alpha, 0 \leqq \alpha \leqq 1\}$. Hence, in general, the set of values of $\mathscr{E}_H k(X, Y)$ when H ranges over a convex family of distributions is not necessarily convex; in fact it has one of the following forms I, $\{-\infty\} \cup I$, $I \cup \{+\infty\}$, $\{-\infty\} \cup I \cup \{+\infty\}$, where I is an interval. (I may be open, semi-open, or closed as well as bounded or unbounded.) We now show in Theorem 2 that if k is quasi-monotone and some further assumptions are satisfied, when H ranges over $\mathscr{H}(F, G)$ the set of values of $\mathscr{E}_H k(X, Y)$ is closed and convex and its supremum and infimum are determined.

The proof of Theorem 2 rests on the following property which is stated separately since it does not require k to be quasi-monotone.

Lemma. *Let X and Y be random variables with distribution functions F and G respectively and joint distribution function H, and let $k(x, y)$ be a Borel measurable, locally bounded function on the plane. If the expectations $\mathscr{E}_{H_+} k(X, Y)$ and $\mathscr{E}_{H_-} k(X, Y)$ exist and at least one of them is finite, then the (possibly unbounded) closed interval with endpoints $\mathscr{E}_{H_+} k(X, Y)$ and $\mathscr{E}_{H_-} k(X, Y)$ belongs to the set of values of $\mathscr{E}_H k(X, Y)$ when H ranges over $\mathscr{H}(F, G)$.*

Proof. According to the discussion preceeding the lemma it suffices to show that if for instance

$$-\infty < \mathscr{E}_{H_+} k(X, Y) < \mathscr{E}_{H_-} k(X, Y) = +\infty$$

there is a sequence $H_n \in \mathscr{H}(F, G)$, $n = 1, 2, \ldots$, such that $\mathscr{E}_{H_n} k(X, Y) \to +\infty$ (all remaining cases can be treated similarly). From the definition of H_+ and H_- we have that under H_+, $(X, Y) \stackrel{d}{=} (F^{-1}(U), G^{-1}(U))$, and under H_-,

$$(X, Y) \stackrel{d}{=} (F^{-1}(U), G^{-1}(1-U)),$$

where U is a uniform random variable on $(0, 1)$ and $F^{-1}(u) = \inf\{t: F(t) \geqq u\}$. Also, by assumption, we have

$$\mathscr{E}_{H_-} k(X, Y) = \mathscr{E} k(F^{-1}(U), G^{-1}(1-U)) = \int_0^1 k(F^{-1}(u), G^{-1}(1-u))\, du = +\infty.$$

Now for each $0 \leqq \alpha \leqq \frac{1}{2}$ define $g_\alpha(u)$ on $[0, 1]$ by

$$g_\alpha(u) = \begin{cases} G^{-1}(1-u) & \text{if } \alpha < u < 1-\alpha \\ G^{-1}(u) & \text{if } 0 \leqq u \leqq \alpha \text{ or } 1-\alpha \leqq u \leqq 1, \end{cases}$$

and let H_α be the distribution function of the pair $(F^{-1}(U), G_\alpha(U))$. For each x we have

$$\begin{aligned}\Pr\{g_\alpha(U)<x\} &= \mathrm{Leb}\{(\alpha, 1-\alpha)\cap(1-G(x), 1]\}\\ &\quad+\mathrm{Leb}\{([0,\alpha]\cup[\alpha-1, 1])\cap[0, G(x))\}=G(x)\end{aligned}$$

and thus $H_\alpha\in\mathscr{H}(F, G)$. We also have

$$\begin{aligned}\mathscr{E}_{H_\alpha}k(X, Y) &= \mathscr{E}k(F^{-1}(U), g_\alpha(U))\\ &= \int_\alpha^{1-\alpha} k(F^{-1}(u), G^{-1}(1-u))\,du+\left(\int_0^\alpha+\int_{1-\alpha}^1\right)k(F^{-1}(u), G^{-1}(u))\,du.\end{aligned}$$

Since for $u\in[\alpha, 1-\alpha]$, $F^{-1}(u)$ and $G^{-1}(1-u)$ are bounded, and since k is locally bounded, it follows that the first integral is finite and

$$\begin{aligned}\lim_{\alpha\downarrow 0}\int_\alpha^{1-\alpha} k(F^{-1}(u), G^{-1}(1-u))\,du &= \int_0^1 k(F^{-1}(u), G^{-1}(1-u))\,du\\ &= \mathscr{E}_{H_-}k(X, Y)=+\infty.\end{aligned}$$

On the other hand, since $\int_0^1 |k(F^{-1}(u), G^{-1}(u))|\,du=\mathscr{E}_{H_+}|k(X, Y)|<+\infty$, we have

$$\lim_{\alpha\downarrow 0}\left(\int_0^\alpha+\int_{1-\alpha}^1\right)k(F^{-1}(u), G^{-1}(u))\,du=0.$$

It follows that $\lim_{\alpha\downarrow 0}\mathscr{E}_{H_\alpha}k(X, Y)=+\infty$, and thus the proof is complete.

Theorem 2. *Let X and Y be random variables with distribution functions F and G respectively and joint distribution function H, and let $k(x, y)$ be quasi-monotone and right continuous. If the expectations $\mathscr{E}_{H_-}k(X, Y)$ and $\mathscr{E}_{H_+}k(X, Y)$ exist (even if infinite valued), then the set of values of $\mathscr{E}_H k(X, Y)$ when H ranges over $\mathscr{H}(F, G)$ is the (possibly unbounded) closed interval $[\mathscr{E}_{H_-}k(X, Y), \mathscr{E}_{H_+}k(X, Y)]$ when either of the following is satisfied:*

(a) *$k(x, y)$ is symmetric and* (i) *holds (in this case $-\infty\leqq\mathscr{E}_{H_-}k(X, Y)\leqq\mathscr{E}_{H_+}k(X, Y)<+\infty$),*

(b) *for some x_0 and y_0,* (ii) *holds and at least one of $\mathscr{E}_{H_+}k(X, Y)$ and $\mathscr{E}_{H_-}k(X, Y)$ is finite.*

Thus under the assumptions of Theorem 2 the infimum and the supremum of $\mathscr{E}_H k(X, Y)$ for $H\in\mathscr{H}(F, G)$ are achieved by H_- and H_+ respectively and they are given by

$$\mathscr{E}_{H_+}k(X, Y)=\int_0^1 k(F^{-1}(u), G^{-1}(u))\,du,$$
$$\mathscr{E}_{H_-}k(X, Y)=\int_0^1 k(F^{-1}(u), G^{-1}(1-u))\,du.$$

Of course if k is quasi-antitone the infimum is achieved by H_+ and the supremum by H_-.

Proof. (a) It is clear from (3) and (5) that for all $H\in\mathscr{H}(F, G)$, $\mathscr{E}_H k(X, Y)$ exists (even if infinite valued) and satisfies

$$-\infty\leqq\mathscr{E}_{H_-}k(X, Y)\leqq\mathscr{E}_H k(X, Y)\leqq\mathscr{E}_{H_+}k(X, Y)<+\infty.$$

If $\mathscr{E}_{H_+}k(X, Y)=-\infty$ then $\mathscr{E}_H k(X, Y)=-\infty$ for all $H\in\mathscr{H}(F, G)$. If

$$-\infty<\mathscr{E}_{H_+}k(X, Y)<+\infty$$

the result follows trivially when $-\infty < \mathscr{E}_{H_-} k(X, Y)$, and from the Lemma when $\mathscr{E}_{H_-} k(X, Y) = -\infty$.

(b) is shown similarly.

It is clear from the discussion in Section 2 that (i) and (ii) in (a) and (b) can be replaced by (i)′ and (ii)′ and the result of Theorem 2 remains valid provided H is restricted to range only over those members of $\mathscr{H}(F, G)$ for which (i)′ and (ii)′ are satisfied (otherwise $\mathscr{E}_H k(X, Y)$ may not exist).

It should be clear that the result of Theorem 2 is no longer true when k is not quasi-monotone (or quasi-antitone), i.e. for general k the closed interval with endpoints $\mathscr{E}_{H_-} k(X, Y)$ and $\mathscr{E}_{H_+} k(X, Y)$ is a proper subset of the set of values of $\mathscr{E}_H k(X, Y)$ when H ranges over $\mathscr{H}(F, G)$. As an example take $k(x, y) = (x-\frac{1}{2})^2 (y-\frac{1}{2})^2$ and F, G to be the uniform distributions on $(0, 1)$. Then under H_+, $X = Y$, and under H_-, $X = -Y$, and thus $\mathscr{E}_{H_+} k(X, Y) = \mathscr{E}_{H_-} k(X, Y) = \mathscr{E}(X-\frac{1}{2})^4 = \frac{1}{80}$. On the other hand when $H(x, y) = F(x) G(y)$ we have $\mathscr{E}_H k(X, Y) = \{\mathscr{E}(X-\frac{1}{2})^2\}^2 = \frac{1}{144}$.

4. Examples and Discussion

Some simple examples of continuous quasi-monotone and quasi-antitone functions are the following. Quasi-monotone functions: $x y$, $(x+y)^2$, $\min(x, y)$, $f(x-y)$ where f is concave and continuous. Quasi-antitone functions: $|x-y|^p$ for $p \geqq 1$, $\max(x, y)$, $f(x-y)$ where f is convex and continuous.

If $k(x, y)$ is absolutely continuous then $\dfrac{\partial^2 k(x, y)}{\partial x \partial y}$ exists a.e. [Leb] and is locally integrable and for all $x \leqq x'$ and $y \leqq y'$,

$$\Delta_{(x, x')}^{(y, y')} k = \int_x^{x'} \int_y^{y'} \frac{\partial^2 k(u, v)}{\partial u \partial v} du\, dv$$

(see Hobson [7]). Hence when $k(x, y)$ is absolutely continuous, it is quasi-monotone if and only if $\partial^2 k(x, y)/\partial x \partial y \geqq 0$ a.e. Thus starting from any non-negative locally Lebesgue integrable function one can generate (absolutely continuous) quasi-monotone functions.

An example of a quasi-antitone function which is not necessarily continuous is $k(x, y) = |f(x) - f(y)|$ where f is nondecreasing, say; when f is right continuous so is k.

Of the two properties required of the function $k(x, y)$ quasi-monotonicity is the crucial one in our method, while right continuity can be somewhat weakened. Of course if $k(x, y)$ is left continuous we can get the same results simply by defining μ by

$$\mu\{[x, x') \times [y, y')\} = \Delta_{(x, x')}^{(y, y')} k.$$

More importantly, if $k(x, y)$ has left and right limits at every point and if its points of discontinuity are located on a countable number of parallels to the axes, then we can obtain the same results provided the points where these parallels cut the axes are not atoms of the marginal distributions F and G (this is of course always satisfied if F and G have densities). This follows from the fact that we can write $k=k_1+k_2$ where k_1 is right continuous and k_2 equals zero for points not on the countable number of parallels to the axes. Then $\mathscr{E}k_2(X, Y)=\int_{R^2} k_2\,dH=0$ since every H in $\mathscr{H}(F, G)$ assigns zero measure to every line parallel to the axes cutting the axis at a point which is not an atom of the corresponding marginal. In this connection it is interesting to note that if $k(x, y)$ is quasi-monotone and if for some x_0 and y_0, $k(x_0, y)$ and $k(x, y_0)$ are of bounded variation in y and x respectively then $k(x, y)$ has left and right limits at every point and its points of discontinuity are located on a countable number of parallels to the axes (for the first part see p. 345 of Hobson [7] and for the second p. 722 of Adams and Clarkson [1]).

A function $k(x, y)$ is said to be of bounded variation on a bounded rectangle $[a, b]\times[c, d]$ if for all m, n and points

$$a=x_0<x_1<\cdots<x_m=b, \qquad c=y_0<y_1<\cdots<y_n=d$$

the sum $\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}|\Delta_{(x_i, x_{i+1})}^{(y_i, y_{i+1})}k|$ is bounded. If k is of bounded variation on every bounded rectangle, it is the difference of two quasi-monotone functions (see p. 718 of Adams and Clarkson [1]) and if it is also right continuous it determines by (1) a (unique σ-finite) signed measure μ. Then under the appropriate integrability assumptions the expressions for $\mathscr{E}k(X, Y)$ given by (3) and (5) and by (7) and (9) remain valid. Since μ can now take both positive and negative values the results of Theorems 1 and 2 are no longer valid. However, one can still obtain weaker results some of which we mention briefly.

(i) If the joint distributions H and H' of (X, Y) and (X', Y') as in Theorem 1 assign full measure to Borel sets which are positive sets of the signed measure μ, then (10) is valid. Of course this means that k is quasi-monotone with respect to H and H', i.e., on sets of full H and H' measure rather than on the entire plane. As an example $k(x, y)=|x-y|^p$, $0<p<1$, is quasi-monotone on the complement of the diagonal of the plane and thus if $\Pr\{X=Y\}=0$ under H and $\Pr\{X'=Y'\}=0$ under H' (i.e. H and H' assign zero probability to the diagonal), (10) is valid. An even simpler case arises when with probability one $X\in A$ and $Y\in B$ where the Borel sets A and B are such that k is quasi-monotone on $A\times B$ (but not necessarily on the entire plane). Then Theorems 1 and 2 remain valid without any further qualifications. As examples, take $k(x, y)=(x+y)^p$, $1\leqq p$, and $X\geqq 0$, $Y\geqq 0$ with probability one, or $k(x, y)=|x-y|^p$, $0<p<1$, and $X\leqq a<b\leqq Y$ with probability one.

(ii) If $\mu=\mu_1-\mu_2$ is the Jordan decomposition of the signed measure μ as a difference of two nonnegative measures μ_1 and μ_2, under appropriate integrability conditions, one can get upper and lower bounds for $\mathscr{E}_H k(X, Y)$, $H\in\mathscr{H}(F, G)$, like those in Theorem 2. For instance for k symmetric (under appropriate integrability conditions) we have that for all H in $\mathscr{H}(F, G)$,

$$\int_{R^2} A_+\,d\mu_2-\int_{R^2} A_-\,d\mu_1\leqq 2\mathscr{E}k(X, Y)-\mathscr{E}k(X, X)-\mathscr{E}k(Y, Y)$$
$$\leqq\int_{R^2} A_-\,d\mu_2-\int_{R^2} A_+\,d\mu_1$$

where A_+, A_- are given by (3) with H replaced by H_+, H_- respectively. However, these upper and lower bounds are not achieved by some H's in $\mathscr{H}(F, G)$ and more important they are not the least upper bound and the greatest lower bound (whose calculations elude us).

5. Discussion of the Literature

We conclude with a few comments on the literature. For $k(x,y)=|x-y|^p$ the expression of $\mathscr{E}k(X,Y)$ given by (3) and (5) has been obtained for $p=2$ by Hoeffding [8] (see also p. 1139 of Lehmann [9]) and by Bártfai [2], for $p=1$ by Vallender [14], and for any $p\geqq 1$ by Dall'Aglio [3] (see also Dall'Aglio [4]). For $k(x,y)=xy$ the bounds of Theorem 2 are given on p. 278 of Hardy, Littlewood and Pólya [6] by the method of rearrangements. This work was done independently of Tchen [13] where the inequality of Theorem 1 is derived for continuous and bounded quasi-monotone functions.

References

1. Adams, C.R., Clarkson, J.A.: Properties of functions $f(x,y)$ of bounded variation. Trans. Amer. Math. Soc. **36**, 711–730 (1934)
2. Bártfai, P.: Über die Entfernung der Irrfahrtswege. Studia Sci. Math. Hungar. **5**, 41–49 (1970)
3. Dall'Aglio, G.: Sugli estremi dei momenti delle funzioni di ripartizione doppia. Ann. Sci. École Norm. Sup. **10**, 35–74 (1956)
4. Dall'Aglio, G.: Fréchet classes and compatibility of distribution functions. Symposia Mathematica **9**, 131–150. New York: Academic Press 1972
5. Fréchet, M.: Sur les tableaux de corrélation dont les marges sont données. Ann. Univ. Lyon, Sect. A, Sér. **3**, 14, 53–77 (1951)
6. Hardy, G.H., Littlewood, J.E., Pólya, G.: Inequalities. Cambridge: Cambridge Univ. Press 1952
7. Hobson, E.W.: The Theory of Functions of a Real Variable and the Theory of Fourier's Series. Washington: Harren Press 1950
8. Hoeffding, W.: Maßstabvariante Korrelationstheorie. Schr. Math. Inst. Univ. Berlin **5**, 181–233 (1940)
9. Lehmann, E.L.: Some concepts of dependence. Ann. Math. Statist. **37**, 1137–1153 (1966)
10. Neumann, J. v.: Functional Operators. Volume I: Measures and Integrals. Princeton: Princeton Univ. Press 1950
11. Pledger, G., Proschan, F.: Stochastic comparisons of random processes, with applications in reliability. J. Appl. Probability **10**, 572–585 (1973)
12. Skorokhod, A.V.: Limit theorems for stochastic processes. Theor. Probability Appl. **1**, 261–290 (1965)
13. Tchen, A.H.T.: Exact inequalities for $\int \phi(x,y)\,dH(x,y)$ when H has given marginals. Manuscript (1975)
14. Vallender, S.S.: Calculation of the Wasserstein distance between probability distributions on the line. Theor. Probability Appl. **18**, 784–786 (1973)
15. Veinott, A.F., Jr.: Oprimal policy in a dynamic, single product, nonstationary inventory model with several demand classes. Operations Res. **13**, 761–778 (1965)

Received February 20, 1976