ON PAIRING OBSERVATIONS FROM A DISTRIBUTION WITH MONOTONE LIKELIHOOD RATIO

By Milton C. Chew, Jr.

Rensselaer Polytechnic Institute

It is assumed that a random sample of size n is taken from a bivariate distribution whose density $f(x, y)$ possesses a Monotone Likelihood Ratio, i.e. for all $x_1 < x_2$ and $y_1 < y_2$, $f(x_1, y_1)f(x_2, y_2) \geq f(x_1, y_2)f(x_2, y_1)$. When the sample is "broken," i.e. when the x- and y-values are received in random relative order, it is desirable to optimally "reconstruct" the original bivariate sample. Optimal properties of the Maximum Likelihood Pairing (MLP) of x- and y-values, obtained by DeGroot, *et al.* in [1], are generalized to the class of distributions defined above, with particular attention given to the trinomial distribution. In addition, one of the main results shown is that in general the MLP is better than random pairing, in that the expected number of correct pairings using the MLP is greater than unity.

1. Introduction and summary. The random matching problem furnishes an interesting exercise in combinatorial probability, and is discussed in many elementary texts such as Feller [2]. In any practical matching situation, however, the matchmaker typically has available information which most likely will increase his chances of a successful match if pairing is done judiciously. As an example of such a situation, consider the problem of "decoding" a set of messages transmitted randomly over a noisy, memoryless channel, or communications link. If the statistics of the channel are known to the receiver, he may be able to decode the received messages with greater accuracy.

Suppose then that X and Y are real, measurable quantities associated with the two types of objects to be matched or paired, and that there are a total of n pairs of these objects under consideration. Suppose also that the (x_i, y_i) pairs of measurements for $i = 1, \cdots, n$ resulting from a perfect match, form a random sample of such pairs in an infinite population, and denote by $f(x, y)$ the joint probability density function (possibly discrete) of the random variables X and Y.

To formalize the matching situation discussed above, we assume that a random sample of size n from the distribution with density $f(\cdot, \cdot)$ is available only in the form $(\mathbf{x}; \mathbf{y}) = (x_1, \cdots, x_n; y_1, \cdots, y_n)$, where the x-values and y-values are in random order relative to each other. Such a sample is termed a broken random sample (see [1]). If we denote by ϕ the permutation of y-values relative to x-values which results in pairing x_i with $y_{\phi(i)}$, $i = 1, \cdots, n$, then it is desirable to find that permutation ϕ^* which results in some manner of optimal pairing of x- and y-values. Two particular optimality criteria are of interest: (i) to maximize the probability of a perfect match among all n pairs of x- and y-values, and (ii) to maximize the expected number of correctly matched pairs among all n pairs of values.

DeGroot, Feder and Goel [1] solve this problem under the assumption that the joint density of X and Y has the form:

$$(1.1) \qquad\qquad f(x, y) = a(x)b(y)e^{xy}, \qquad\qquad \text{for } (x, y) \in R^2.$$

Actually, in [1] it is assumed that sets of measurements T and U of dimension r and s respectively can be taken from the two types of objects to be paired, and that the joint $r + s$-dimensional distribution is specified by the joint density:

$$g(t, u) = \alpha(t)\beta(u)e^{\gamma(t)\delta(u)}, \qquad \text{for all } t \in R^r \text{ and } u \in R^s,$$

where α, β, γ and δ are arbitrary real-valued functions. However, this is essentially equivalent to the assumption that there exist real functions γ and δ such that the random variables $X = \gamma(T)$ and $Y = \delta(U)$ have a joint density given by (1.1). Hence, in practice, one is required to reduce the two vector measurements T and U to real measurements prior to matching the sampled vectors $t_1, \cdots, t_n$ with $u_1, \cdots, u_n$.

In [1], the authors demonstrate that the maximum likelihood solution is to pair the x- and y-values in similar order, i.e., if $x_1 \leqq \cdots \leqq x_n$ and $y_1 \leqq \cdots \leqq y_n$, then x_i should be paired with y_i, $i = 1, \cdots, n$. If we call this solution the Maximum Likelihood Pairing (MLP), the authors then show that the MLP maximizes the posterior probability of a perfect match, given the values of the broken random sample $(\mathbf{x}; \mathbf{y})$, and obtain sufficient conditions for an MLP to maximize the expected number of correct pairings, given the data.

The purpose of this paper is (i) to extend the results of DeGroot, *et al.* to a slightly more general class of bivariate distributions—those whose densities possess a Monotone Likelihood Ratio (MLR); and (ii) to establish an additional property of the MLP.

Section 2 deals with the extension of the results in [1], and in Section 3 we consider an example of matching a broken random sample from a trinomial distribution, whose density possesses an MLR, but does not satisfy (1.1). In Section 4, it is shown that if $f(\cdot, \cdot)$ possesses an MLR, then the expected number of correct pairings using an MLP is at least unity for all values $(\mathbf{x}; \mathbf{y})$ of a broken random sample. Since the expected number of correct pairings under random matching is one, the MLP can never do worse than random pairing, hence complementing the optimality results of [1] and Section 2. It is also argued that in order to always do better than random pairing, the matchmaker should pair x_1 with y_1, and x_n with y_n.

2. Generalizations of previous results. Suppose that $(\mathbf{x}; \mathbf{y}) = (x_1, \cdots, x_n; y_1, \cdots, y_n)$ is a broken random sample from a bivariate distribution with density $f(\cdot, \cdot)$. Without loss of generality, we assume that $x_1 \leqq \cdots \leqq x_n$ and $y_1 \leqq \cdots \leqq y_n$. We shall say that $f(\cdot, \cdot)$ possesses a Monotone Likelihood Ratio (MLR) if for all $x_1 < x_2$ and $y_1 < y_2$,

$$(2.1) \qquad\qquad f(x_1, y_1)f(x_2, y_2) \geqq f(x_1, y_2)f(x_2, y_1).$$

Note that all densities satisfying (1.1) will possess an MLR as defined above. Strictly speaking, f will possess a non-decreasing MLR if it satisfies (2.1), and we could similarly define f to have a non-increasing MLR if (2.1) holds whenever $x_1 \leq x_2$ and $y_1 \geq y_2$. An example of the latter is the trinomial distribution which we examine more closely in Section 3. Let Φ be the set of all permutations of $(1, \cdots, n)$. A pairing of x- and y-values according to $\phi \in \Phi$ is identified by pairing x_i with $y_{\phi(i)}$, $i = 1, \cdots, n$. The likelihood function associated with a permutation ϕ is then the joint density function of $(x_i, y_{\phi(i)})$, $i = 1, \cdots, n$, i.e.,

$$(2.2) \qquad L(\phi; \mathbf{x}, \mathbf{y}) = \prod_{i=1}^{n} f(x_i, y_{\phi(i)}) \,.$$

Since f possesses an MLR, the maximum likelihood pairing (MLP), which maximizes (2.2) among all $\phi \in \Phi$ is the natural order, i.e., that permutation which pairs x_i with y_i, $i = 1, \cdots, n$. To see this, consider a permutation ϕ for which $y_{\phi(j)} < y_{\phi(i)}$ for some $x_i < x_j$. Then by (2.1), $f(x_i, y_{\phi(j)}) f(x_j, y_{\phi(i)}) \geq f(x_i, y_{\phi(i)}) \times f(x_j, y_{\phi(j)})$. Thus, $L(\psi; \mathbf{x}, \mathbf{y}) \geq L(\phi; \mathbf{x}, \mathbf{y})$, where $\psi \in \Phi$ is such that $\psi(i) = \phi(j)$, $\psi(j) = \phi(i)$, and $\psi(k) = \phi(k)$ for all other integers k. Since this process of the pairwise interchange of ϕ-coordinates not in natural order could be continued, we conclude that $L(\phi^*; \mathbf{x}, \mathbf{y}) = \max_\phi L(\phi; \mathbf{x}, \mathbf{y})$, where $\phi^* = (1, \cdots, n)$. Further, since the x- and y-values of the broken random sample are in relative random order, each pairing $\phi \in \Phi$ is equally likely a priori to yield the original bivariate sample. Therefore, the posterior probability of a perfect match using ϕ, given the values of the broken random sample $(\mathbf{x}; \mathbf{y})$, is

$$(2.3) \qquad p(\phi) = L(\phi; \mathbf{x}; \mathbf{y}) / L(\mathbf{x}, \mathbf{y}) \,,$$

where $L(\mathbf{x}, \mathbf{y}) = \sum_{\phi \in \Phi} L(\phi; \mathbf{x}, \mathbf{y})$, and is maximized by the MLP ϕ^*. This and other results demonstrating the optimality of ϕ^*, are established by similar arguments in [1] for distribution whose densities satisfy (1.1). The class of distributions for which these results hold can be enlarged, as above, to include those which possess an MLR. In fact, the extension is trivial in that it essentially only requires the replacement of the factor $\exp(x_i y_k)$ by $f(x_i, y_k)$ wherever it appears in the proofs given in [1].

The extent to which this "replacement" can occur is limited, however, when one attempts to maximize the expected number of correct pairings. Specifically, if we let $N(\phi, \zeta)$ denote the number of common (x, y) pairs for any permutations ϕ and ζ in Φ, the expected number of correct pairings when x_i is paired with $y_{\phi(i)}$, $i = 1, \cdots, n$, may be written as:

$$(2.4) \qquad M(\phi) = \sum_{\zeta \in \Phi} N(\phi, \zeta) p(\zeta) \,.$$

If ϕ is any permutation for which $y_{\phi(j)} < y_{\phi(i)}$ when $x_i < x_j$ for some i and j, then one desires to show $M(\psi) \geq M(\phi)$, where ψ is the permutation such that $\psi(i) = \phi(j)$, $\psi(j) = \phi(i)$, and $\psi(k) = \phi(k)$ for all other integers k. This would then establish that the MLP ϕ^* maximizes the expression in (2.4). Unfortunately this does not hold in general, and DeGroot, et al. ([1]), focus their attention on finding conditions for which $M(\psi) \geq M(\phi)$ holds for all $\phi \neq \phi^*$. In their arguments,

the set of permutations Φ is partitioned into mutually exclusive subsets according as $\zeta \in \Phi$ is such that $\mathbf{y}_\zeta = (y_{\zeta(1)}, \cdots, y_{\zeta(n)})$ disagrees with $\mathbf{y}_\phi$ in neither, both, or one of the ith or jth coordinates. In fact, the latter set is further partitioned according as the remaining, or unmatched, ith or jth coordinate of $\mathbf{y}_\zeta$ is inside or outside the interval determined by $y_{\phi(j)}$ and $y_{\phi(i)}$. A one-to-one correspondence exists between appropriate pairs of subsets of Φ, and it can be shown (see [1]) by using the MLR property of $f(\cdot, \cdot)$ that $M(\phi; \Phi^*) \geqq M(\phi; \Phi^*)$, where $M(\phi; \Phi^*) = \sum_{\zeta \in \Phi^*} N(\phi, \zeta) p(\zeta)$ and Φ^* includes all permutations except those last cited above; i.e., $\zeta \in \Phi - \Phi^*$ if only one of the ith or jth coordinates of $\mathbf{y}_\zeta$ differs with $y_{\phi(i)}$ or $y_{\phi(j)}$, and is either less than $y_{\phi(j)}$ or greater than $y_{\phi(i)}$. As in [1], when f has MLR, it can be shown that a condition sufficient for $M(\phi; \Phi - \Phi^*) \geqq M(\phi; \Phi - \Phi^*)$ to hold is that

$$(2.5) \qquad A = f(x_i, y_q)g(x_h, x_j; y_{\phi(j)}, y_{\phi(i)}) + f(x_j, y_q)g(x_i, x_h; y_{\phi(j)}, y_{\phi(i)}) \geqq 0 \,,$$

for all pairs (x_h, y_q) such that x_h differs from x_i and x_j, and $y_q < y_{\phi(j)}$ or $> y_{\phi(i)}$. Here, we have set $g(a, b; c, d) = f(a, c)f(b, d) - f(a, d)f(b, c)$ for all a, b, c and d. It is difficult to simplify this condition, as done in [1] for distributions satisfying (1.1), unless the form of the density is known. We can note, however, that since f has MLR, the g-function in both terms of A is always nonnegative whenever $x_i < x_h < x_j$. Summarizing then, we have the following analogue of Theorem 1 of [1]:

THEOREM 1. *Let $f(\cdot, \cdot)$ have* MLR, *and suppose that $\phi \in \Phi$ is a permutation such that $y_{\phi(i)} > y_{\phi(j)}$ when $x_i < x_j$ for some $1 \leqq i < j \leqq n$. If ψ is the permutation defined by $\psi(i) = \phi(j)$, $\psi(j) = \phi(i)$ and $\psi(k) = \phi(k)$ for all other values of k, then $M(\psi) \geqq M(\phi)$, whenever in (2.5), $A \geqq 0$ for all (h, q) pairs such that $x_h \neq [x_i, x_j]$ and $y_q \neq [y_{\phi(j)}, y_{\phi(i)}]$.*

The solution of the problem of maximizing the expected number of correct pairings, for the enlarged class of distributions with MLR, appears to have marginal value at best, in light of the complicated condition in the above theorem. However, when the form of the joint density $f(\cdot, \cdot)$ is known, it may be possible to reduce the complexity of condition (2.5), and even replace it by a remarkably simple sufficient condition that is applicable to all permutations ϕ. In [1] for instance, the authors show that if $f(\cdot, \cdot)$ satisfies (1.1), a sufficient condition for the MLP ϕ^* to maximize $M(\phi)$ is that

$$(2.6) \qquad (x_n - x_1)(y_n - y_1) \leqq 1 \,.$$

In particular, when the y-values are linearly transformed so that (1.1) accommodates the bivariate normal density, then ϕ^* maximizes $M(\phi)$ whenever

$$(2.7) \qquad (x_n - x_1)(y_n - y_1) \leqq (1 - \rho^2)\sigma_1\sigma_2/\rho \,,$$

where σ_1 and σ_2 are the standard deviations of X and Y respectively, and ρ, assumed positive, is the correlation between X and Y.

In the next section, we examine more closely the situation when the joint distribution of X and Y is a trinomial distribution. This is an example of a distribution possessing a (non-increasing) MLR, and such that the joint density of X and Y is not of the type given in (1.1).

3. Sampling from a trinomial distribution. Suppose now that $(\mathbf{x}; \mathbf{y}) = (x_1, \cdots, x_n; y_1, \cdots, y_n)$ is a broken random sample from a trinomial distribution with discrete probability density given by:

$$(3.1) \qquad f(x, y) = \binom{N}{x}\binom{N-x}{y}p^x q^y (1 - p - q)^{N-x-y} ,$$

for x and y nonnegative integers such that $x + y \leq N$. The density (3.1) possesses a non-increasing MLR, since to say that $f(x_1, y_1)f(x_2, y_2) \geq f(x_1, y_2)f(x_2, y_1)$ whenever $x_1 \leq x_2$ and $y_1 \geq y_2$, is equivalent to saying that

$$\binom{N-x_1-y_2}{x_2-x_1} \geq \binom{N-x_1-y_1}{x_2-x_1} .$$

The latter holds, however, since the binomial coefficient $\binom{a}{b}$ is increasing in a.

For convenience, let us re-index the values in the broken sample $(\mathbf{x}; \mathbf{y})$ so that $x_1 \leq \cdots \leq x_n$ and $y_1 \geq \cdots \geq y_n$. In this instance, the MLP ϕ^* is again the natural order, since the trinomial density (3.1) has non-increasing MLR. Thus, the maximum likelihood solution is to pair the x- and y-values in dissimilar order. As a result of the reordering, the theorem of the previous section is now applicable to the distribution in (3.1), and we wish to examine the condition $A \geq 0$ in (2.5) in more detail. It should be pointed out, however, that in dealing with the trinomial distribution, the possibility exists that some pairings of x- and y-values may yield 'inadmissible' samples, i.e. it is possible that a permutation ϕ may be such that for some i, $x_i + y_{\phi(i)} > N$ in which case the likelihood $L(\phi; \mathbf{x}, \mathbf{y}) = 0$. Such permutations should not be excluded from consideration since it is conceivable that the expected number of correct pairings is maximized by such a ϕ. Note though that by admitting all possible pairings of x- and y-values, i.e. by allowing a pair (x, y) for which $f(x, y) = 0$, the MLR property of $f(\cdot, \cdot)$ is not disturbed, and the results of the previous section remain applicable.

On examining the condition $A \geq 0$ in (2.5), it can easily be seen that if any (x, y) pair involved in (2.5) is not admissible, i.e. $x + y > N$ and hence $f(x, y) = 0$, then $A \geq 0$ whenever $x_h < x_i$ and $y_q > y_{\phi(j)}$, or $x_h > x_j$ and $y_q < y_{\phi(i)}$. If all pairs are admissible, then letting $r = \phi(j)$ and $s = \phi(i)$, (2.5) reduces to

$$(3.2) \qquad \binom{N-x_i-y_s}{x_j-x_i}[\binom{N-x_h-y_s}{y_r-y_s} - \binom{N-x_j-y_s}{y_r-y_s}] - \binom{N-x_i-y_q}{x_j-x_i}[\binom{N-x_h-y_s}{y_r-y_s} - \binom{N-x_i-y_s}{y_r-y_s}] \geq 0 ,$$

which holds whenever $x_h < x_i$ and $y_q > y_r$, or $x_h > x_j$ and $y_q < y_s$. Thus, we have the following corollary to Theorem 1:

COROLLARY 1. *Let* $(\mathbf{x}; \mathbf{y})$ *be a broken random sample from the trinomial distribution with density given in* (3.1). *If* ϕ *is any pairing such that* $y_{\phi(i)} < y_{\phi(j)}$ *for some* i *and* j *such that* $x_i < x_j$, *and* ψ *is the permutation defined by* $\psi(i) = \psi(j) = r$,

$\phi(j) = \phi(i) = s$, and $\phi(k) = \phi(k)$ *for all other values of* k, *then* $M(\psi) \geqq M(\phi)$, *whenever* (2.5) *holds for all pairs* (x_h, y_q) *such that either* $x_h < x_i$ *and* $y_q < y_s$, *or* $x_h > x_j$ *and* $y_q > y_r$.

This result now enables us to state the following analogue of Theorem 2 of [1]:

THEOREM 2. *Let* $(\mathbf{x}; \mathbf{y})$ *be a broken random sample from the trinomial distribution. Then there exists a pairing* ϕ *which maximizes the expected number of correct pairs such that* $\phi(1) = 1$ *and* $\phi(n) = n$.

The proof is precisely as in [1]. If ϕ is any permutation such that $y_{\phi(1)} < y_1$ and $\phi(j) = 1$ for some j such that $x_j > x_1$, let ψ be that permutation as defined in the above corollary with $i = 1$. Then the conditions of the corollary are vacuously satisfied with $M(\psi) \geqq M(\phi)$. Similarly, if ϕ is such that $y_{\phi(n)} > y_n$ and $\phi(i) = n$ for some i such that $x_i < x_n$, then with $j = n$, the ψ in the corollary again satisfies $M(\psi) \geqq M(\phi)$.

Let us now examine (3.2) in more detail in an attempt to find simpler sufficient conditions for $M(\psi) \geqq M(\phi)$. For (3.2) to be applicable, we must require the permutation ϕ to assign only admissible pairs, i.e. $x_i + y_{\phi(i)} \leqq N$ for all $i = 1, \cdots, n$.

Case I. $x_h < x_i$ *and* $y_q < y_s$.

If $i = 1$ or $s = n$, (3.2) is vacuously satisfied so it is assumed that $x_i > x_1$ and $y_s > y_n$. Now, as a function of x_h and y_q, the left side of (3.2) is seen to be increasing in each variable. Hence if we replace them by their smallest possible values, x_1 and y_n respectively, and denote the resulting quantity on the left of (3.2) by $C_1(x_j; x_i)$, it follows that (3.2) will be satisfied for all pairs (x_h, y_q) such that $x_h < x_i$ and $y_q < y_s$ if $C_1(x + 1; x) \geqq 0$ for all $x = x_i, \cdots, x_j - 1$. But $C_1(x + 1; x) \geqq 0$ reduces to

$$(3.3) \qquad (y_r - y_n) - (y_s - y_n)R_1(x) \geqq 0 ,$$

where

$$R_1(x) = \binom{N - x_1 - y_s}{y_r - y_s} / \binom{N - x - y_s}{y_r - y_s} .$$

The quantity on the left in (3.3) however, is decreasing in x, and thus holds for all required x-values if it holds for the largest such value, $x_j - 1$. Denote the resulting quantity by $D_1(y_r; y_s)$, and note that $D_1(y_r; y_s) \geqq 0$ if $D_1(y + 1; y) \geqq 0$ for all $y = y_s, \cdots, y_r - 1$. Since $D_1(y + 1; y) \geqq 0$ reduces to

$$(3.4) \qquad N - x_j - y + 1 - (y - y_n)(x_j - x_1 - 1) \geqq 0 ,$$

we again observe that since the left side of (3.4) is decreasing in y, it suffices to have it hold for the largest required y-value, $y_r - 1$, yielding

$$(3.5) \qquad N - x_1 - y_n + 1 \geqq (y_r - y_n)(x_j - x_1) .$$

Case II. $x_h > x_j$ *and* $y_q > y_r$.

If $j = n$ or $r = 1$, (3.2) is again vacuously satisfied, so we assume $x_j < x_n$ and

$y_r < y_1$. As a function of x_h and y_q, the left side of (3.2) is now decreasing in each variable, so it suffices to require (3.2) to hold when x_h and y_q are replaced by their largest possible values. If we denote the left side of (3.2) by $C_2(x_j; x_i)$ when x_h and y_q are replaced by x_n and y_1 respectively, then (3.2) will be satisfied for all pairs (x_h, y_q) such that $x_h > x_j$ and $y_q > y_r$, if $C_2(x + 1; x) \geqq 0$ for all $x = x_i, \cdots, x_j - 1$. As in the previous case, $C_2(x + 1; x) \geqq 0$ reduces to

$$(3.3') \qquad (y_1 - y_s)R_2(x) - (y_1 - y_r) \geqq 0,$$

where

$$R_2(x) = \binom{N - x_n - y_s}{y_r - y_s} / \binom{N - x - y_s}{y_r - y_s}.$$

But the left side of (3.3') is increasing in x, so it suffices for (3.3') to hold for $x = x_i$. When $x = x_i$, write (3.3') as $D_2(y_r; y_s) \geqq 0$. It then suffices to require $D_2(y + 1; y) \geqq 0$ for all $y = y_s, \cdots, y_r - 1$, which reduces to

$$(3.4') \qquad N - x_i - y - (y_1 - y)(x_n - x_i) \geqq 0.$$

The left side is increasing in y, and is satisfied for all $y = y_s, \cdots, y_r - 1$, if it holds for $y = y_s$, yielding

$$(3.5') \qquad N - x_i - y_s \geqq (y_1 - y_s)(x_n - x_i).$$

Summarizing, then, we have the following theorem.

THEOREM 3. *Let the permutations ϕ and ψ be as in Corollary 1 above, and suppose all pairs of ϕ are admissible. If $i = 1$ or $s = n$ or (3.5) is satisfied, and if $j = n$ or $r = 1$ or (3.5') is satisfied, then $M(\psi) \geqq M(\phi)$.*

As a consequence of this theorem it is easy now to obtain a sufficient condition for optimality of the MLP ϕ^* similar to (2.6).

THEOREM 4. *If $(\mathbf{x}; \mathbf{y})$ is a broken random sample from the trinomial distribution given by (3.1), and if*

$$(3.6) \qquad (x_n - x_1)(y_1 - y_n) \leqq N - x_1 - y_n,$$

then the expected number of correct pairs is maximized by the MLP ϕ^*, i.e. the permutation that pairs x- and y-values in dissimilar order: x_i with y_i for $i = 1, \cdots, n$.

To prove the theorem, let ϕ be any permutation, all of whose (x, y) pairs are admissible, and such that $y_{\phi(i)} < y_{\phi(j)}$ when $x_i < x_j$ for some i and j. If (3.6) is satisfied, then it is easily seen that both (3.5) and (3.5') are satisfied. Hence $M(\phi^*) \geqq M(\phi)$ for any such ϕ. It only remains to show that (3.6) admits no permutations ϕ which assign at least one inadmissible pair that may be better than ϕ^*. Suppose then that ϕ' is a permutation that assigns an inadmissible pair (x, y), i.e. $x + y > N$. It is easily shown by applying (3.6) that $(y - y_n - 1)(x - x_1 - 1) \leqq 0$. Since necessarily $x > x_1$ and $y > y_n$, the only *possible* (x, y) pairs are $(x_1 + 1, y_1)$ or $(x_n, y_n + 1)$. However, from Theorem 2 we know there exists a ϕ which pairs x_1 with y_1 and x_n with y_n such that $M(\phi) \geqq M(\phi')$. If ϕ satisfies (3.6) it is necessarily admissible and hence $M(\phi^*) \geqq M(\phi) \geqq M(\phi')$, establishing the theorem.

Clearly, condition (3.6) is not vacuous, but there is some question as to the necessity of (3.6). That is, it may be possible that the MLP ϕ^* *always* maximizes $M(\phi)$. This is not the case, at least for sufficiently large N, however, as a few isolated examples have shown. Suppose $n = 4$, $\phi = (1, 2, 3, 4)$ and $\psi = (1, 3, 2, 4)$. Then $M(\psi) > M(\phi)$ in the following cases: (i) $N = 1800$, $\mathbf{x} = (600, 601, 602, 700)$, $\mathbf{y} = (600, 452, 451, 450)$; (ii) $N = 200$, $\mathbf{x} = (65, 66, 67, 70)$, $\mathbf{y} = (65, 42, 41, 40)$; (iii) $N = 100$, $\mathbf{x} = (35, 36, 37, 50)$, $\mathbf{y} = (35, 12, 11, 10)$. Presumably, more serious computational effort would yield ranges of data-values for which $M(\psi) > M(\phi)$ for additional values of N. As might be expected, it becomes more difficult (at least when $n = 4$) to find such examples as N takes on smaller values. In fact, if $N = 1$, then in a trivial sense ϕ cannot be exceeded, and it might be of some interest to know how large N must be before it is possible to have $M(\psi) > M(\phi)$.

It is interesting to compare (3.6) with the sufficient condition (2.7) for the bivariate normal distribution, when N is large. For (2.7) to be applicable we require its extension to the case when $\rho < 0$, which is (2.7) with the order of the y-values reversed and ρ replaced by $|\rho|$. The MLP ϕ^* is then the dissimilarly ordered pairing of x- and y-values. Expressing ρ, σ_1 and σ_2 in terms of N, p and q, the appropriate version of (2.7) then becomes

$$(3.7) \qquad\qquad (x_n - x_1)(y_1 - y_n) \leqq N(1 - p - q) \,,$$

which is similar to (3.6) except Np and Nq occur on the right instead of x_1 and y_n. Since x_1 and y_n underestimate Np and Nq, (3.7) is a more 'restrictive' condition, no doubt a consequence of approximating the trinomial by the bivariate normal.

4. MLP is always better than random pairing. In this section we again assume $(\mathbf{x}; \mathbf{y}) = (x_1, \cdots, x_n; y_1, \cdots, y_n)$ is a broken random sample from a bivariate distribution whose density $f(\cdot, \cdot)$ possesses a (non-decreasing) MLR as defined by (2.1). However, we will restrict attention to those samples whose x- and y-values are distinct, i.e. $x_1 < x_2 < \cdots < x_n$ and $y_1 < y_2 < \cdots < y_n$. This is essentially no restriction if the distribution $f(\cdot, \cdot)$ is continuous. Our interest here is to compare the expected number of correct pairings under the MLP ϕ^*—the natural order—with that under random pairing. If the two sets of n quantities to be paired are all distinct, the expected number of correct pairs when they are randomly matched is unity, but is otherwise larger depending of course on the number of non-distinct values among both $\mathbf{x} = (x_1, \cdots, x_n)$ and $\mathbf{y} = (y_1, \cdots, y_n)$. Even though samples from the trinomial distribution, for instance, may be excluded from consideration, the complexity involved by accommodating repeated values in a sample appears at present to overshadow the benefit to be derived. An important advantage in assuming distinct sample values is that now the permutations ϕ completely distinguish between the possible pairings of x- and y-values. We shall exploit this feature toward establishing the following theorem.

THEOREM 5. *If* (x; y) *is a broken random sample from* $f(\cdot, \cdot)$, *possessing* MLR, *and is such that* $x_1 < \cdots < x_n$ *and* $y_1 < \cdots < y_n$, *then the expected number of correct pairings using the* MLP ϕ^* *is at least unity, i.e.*

$$(4.1) \qquad\qquad\qquad M(\phi^*) \geqq 1 .$$

The proof of the theorem will be accomplished by establishing a certain correspondence between sets of permutations. Recalling that Φ is the set of all permutations of $\phi^* = (1, \cdots, n)$, and defining Φ_k to be that subset of Φ whose members have exactly k undisplaced coordinates for $k = 0, 1, \cdots, n$, then $M(\phi^*)$ may be expressed in the following two ways:

$$(4.2) \qquad M(\phi^*) = \sum_{k=1}^{n} \sum_{\phi} \{p(\psi) : \psi(k) = k\} = \sum_{k=1}^{n} k \sum_{\phi} \{p(\psi) : \psi \in \Phi_k\} ,$$

since now the number of common (x, y) pairs under any two permutations ϕ and ψ is *equal* to the number of common coordinates between ϕ and ψ. Both expressions contain an equal number of terms—counting those with coefficient k as k terms in the second expression—and it is easily seen from the first that the number of terms is $n!$, the size of Φ. Since the posterior probabilities $p(\psi)$ sum to one over Φ, both $M(\phi^*)$ and 1 can be expressed as sums of equal numbers of terms $p(\psi)$, i.e. (4.1) is equivalent to

$$\sum_{k=1}^{n} k \sum_{\phi} \{p(\psi) : \psi \in \Phi_k\} \geqq \sum_{\phi \in \Phi} p(\psi) .$$

By cancelling like terms on both sides of the above inequality, we obtain

$$(4.3) \qquad\qquad \sum_{k=2}^{n} (k - 1) \sum_{\phi} \{p(\psi) : \psi \in \Phi_k\} \geqq \sum_{\psi \in \Phi_0} p(\psi)$$

since the sets Φ_k are mutually exclusive and $\Phi = \bigcup_{k=0}^{n} \Phi_k$. If we denote by $r \cdot \Phi'$ a set of permutations that contains exactly r copies of each member of Φ', then the sum on the left of (4.3) could be represented as a sum of $p(\psi)$ over the set $\Phi^* = \bigcup_{k=2}^{n} (k - 1) \cdot \Phi_k$. Since Φ^* and Φ_0 have the same number of members, (4.3)—and consequently (4.1)—will be established if we can define a *positive correspondence* between Φ^* and Φ_0, i.e. if for each $\phi \in \Phi^*$ there is a distinct $\psi \in \Phi_0$ such that $p(\phi) \geqq p(\psi)$.

We now define a partial ordering on the set Φ as follows: if $\phi \in \Phi$ is such that for some $1 \leqq i < j \leqq n$, $\phi(i) > \phi(j)$, and $\psi \in \Phi$ is such that $\psi(i) = \phi(j)$, $\psi(j) = \phi(i)$, and $\psi(k) = \phi(k)$ for all other integers k, we shall say $\psi > \phi$. In general, we shall say $\psi > \phi$ if there exist permutations $\phi_1, \cdots, \phi_m$ such that $\phi_{i+1} > \phi_i$ as above, and $\phi_m = \psi$, $\phi_1 = \phi$. In other words, $\psi > \phi$ if ψ is obtained from ϕ by a sequence of pairwise interchanges of ϕ-coordinates which are in decreasing order. It is now clear from the definition of MLR, that if $\psi > \phi$, $p(\psi) \geqq p(\phi)$. Thus, to establish a positive correspondence between Φ^* and Φ_0, it suffices to show that for every $\psi \in \Phi^*$ there exists a distinct $\phi \in \Phi_0$ such that $\psi > \phi$. When this is the case, we shall say $\Phi^* > \Phi_0$.

In the sequel we will occasionally be required to know the size of the sets Φ_k and their combinations. If $D_{n,k}$ denotes the number of permutations belonging to Φ_k, then $\{D_{n,k}\}$ are the well-known recontres numbers, and according to

Riordan ([4], Chapter 3), the following relationships hold where $D_n \equiv D_{n,0}$ and $D_0 = 1$:

$$\text{(a)} \quad D_{n,k} = \binom{n}{k}D_{n-k}, \qquad k = 0, 1, \cdots, n;$$

(4.4) $$\text{(b)} \quad D_n = nD_{n-1} + (-1)^n = (n-1)(D_{n-1} + D_{n-2}),$$

$$\text{(c)} \quad \sum_{k=0}^{n} D_{n,k} = n!.$$

From (4.4a) and (4.4c) it is easily shown that

$$\sum_{k=2}^{n} (k-1)D_{n,k} = D_n,$$

clarifying that Φ^* and Φ_0 are in one-one correspondence.

Since we shall establish Theorem 5 by an induction argument on n, in what follows a set Φ' of permutations of $(1, \cdots, n)$ will occasionally be expressed as $\Phi'(n)$ to avoid possible confusion. Before proceeding it should also be remarked at this point that a positive correspondence between two sets of permutations of $(1, \cdots, n)$ as defined above, is unaffected by replacing $(1, \cdots, n)$ by any increasing sequence of n integers $(i_1, \cdots, i_n)$. That is, if in the first correspondence $\psi > \phi$, then $(i_{\psi(1)}, \cdots, i_{\psi(n)}) > (i_{\phi(1)}, \cdots, i_{\phi(n)})$ and conversely, establishing a positive correspondence in the second situation.

Consider now the situation when $n = 2$ or $n = 3$. Then $\Phi^*(2) = \{(1, 2)\}$, $\Phi_0(2) = \{(2, 1)\}$, $\Phi^*(3) = \{1, 2, 3), (1, 2, 3)\}$ and $\Phi_0(3) = \{(2, 3, 1), (3, 1, 2)\}$, so it is clear that $\Phi^* > \Phi_0$ in both cases. Suppose then that $\Phi^*(m) > \Phi_0(m)$, for $m = 2, \cdots, n$; i.e., there exists a positive correspondence between Φ^* and Φ_0 for all $m \leq n$. Our problem then is to show that $\Phi^*(n+1) > \Phi_0(n+1)$.

Let $\Phi^*(n+1) = \Phi_1^*(n+1) \cup \Phi_2^*(n+1)$, where Φ_2^* contains all $\phi \in \Phi^*$ such that $\phi(n+1) = n+1$, and Φ_1^* contains all those permutations which remain, i.e. those ϕ such that $\phi(i) = n+1$ for some $i \leq n$. We first consider the set Φ_1^*, partitioning it into mutually exclusive and exhaustive subsets as follows: *if ϕ_1 and ϕ_2 belong to Φ_1^*, they will be assigned to the same subset if the cycles which contain $n+1$ in ϕ_1 and ϕ_2 respectively are identical.* A cycle of a permutation ϕ is simply a subset of the integers (in this case $1, \cdots, n+1$) which is closed under the operation which transforms the natural order ϕ^* to ϕ, i.e. in the notation of Riordan [4], if $(i_1 \cdots i_m)$ is a cycle of ϕ, then in the process of going from ϕ^* to ϕ, $i_1 \to \cdots \to i_m \to i_1$, where $\to$ means "is replaced by". For example, if ϕ has k coordinates in common with ϕ^*, each of these coordinates forms a cycle of length $m = 1$. Also note that distinct cycles associated with the same permutation contain no common elements, and that each permutation is uniquely determined by the specification of its cycles.

Observe now that if $\phi \in \Phi_1^*$, then the length m of the ϕ-cycle containing $n+1$ satisfies $2 \leq m \leq n-1$, since $\phi(n+1) < n+1$ and since ϕ has at least two undisplaced coordinates, i.e. at least two cycles of length 1 each. If the members of a given subset of Φ_1^* as defined above, have their common cycle containing $n+1$ of length m, then that subset contains D_{n+1-m} members, and in fact can be identified with $\Phi^*(n+1-m)$. This is true since all members of the subset

possess at least two but no more than $n + 1 - m$ undisplaced coordinates, all of which are exclusive of the common cycle. From the induction hypothesis, however, we have $\Phi^*(n + 1 - m) > \Phi_0(n + 1 - m)$, since $2 \leq n + 1 - m \leq n - 1$. We have thus established a positive correspondence between each subset of Φ_1^* defined above with the subset of all permutations in Φ_0 whose members possess the same cycle containing $n + 1$ as do all members of the given subset of Φ_1^*. Since the cycle containing $n + 1$ in each member of Φ_0 must have length at least two the only permutations remaining in Φ_0 are those having a single cycle of length $n + 1$, since a cycle of length n implies a cycle of length 1 is present—an impossibility for members of Φ_0.

Let us now consider the subset $\Phi_2^*(n + 1)$ of $\Phi^*(n + 1)$. Since $\Phi_2^*(n + 1)$ contains all $\phi \in \Phi^*(n + 1)$ for which $\phi(n + 1) = n + 1$, we may write $\Phi_2^*(n + 1) = \bigcup_{k=1}^n k \cdot \Phi_k(n) \times \{n + 1\}$, or as follows:

$$(4.5) \qquad \Phi_2^*(n + 1) = \bigcup_{k=1}^n \Phi_k(n) \times \{n + 1\} \cup \Phi^*(n) \times \{n + 1\},$$

since $\Phi^*(n) = \bigcup_{k=2}^n (k - 1) \cdot \Phi_k(n)$. But from the induction hypothesis $\Phi^*(n) > \Phi_0(n)$, so that $\Phi_2^*(n + 1) > \bigcup_{k=0}^n \Phi_k(n) \times \{n + 1\} = \Phi(n) \times \{n + 1\}$.

We now establish a positive correspondence between $\Phi(n) \times \{n + 1\}$ and that subset of $\Phi_0(n + 1)$ whose members contain a single cycle of length $n + 1$. The correspondence will be defined as follows: *assign to each $\phi \in \Phi(n) \times \{n + 1\}$ that $\psi \in \Phi_0(n + 1)$ obtained from ϕ by displacing successively the right-most members of all cycles in ϕ by $\phi(n + 1) = n + 1$.* That is, if ϕ has N cycles—in addition to the 1-cycle $(n + 1)$—and $q_1 > \cdots > q_N$ are their largest members, then ψ is that permutation for which $\psi(n + 1) = \phi(q_1)$, $\psi(q_i) = \phi(q_{i+1})$, $i = 1, \cdots, N - 1$, $\psi(q_N) = n + 1$, and $\psi(j) = \phi(j)$ for all other integers j. This procedure effectively links the cycles of ϕ together, so that $\psi \in \Phi_0(n + 1)$ and possesses a single cycle of length $n + 1$. It is also clear that $\psi < \phi$ since at each stage of the procedure $n + 1$ is displacing a lesser integer in a lesser position. It is not immediately clear, however, that each $\phi \in \Phi(n) \times \{n + 1\}$ is assigned a distinct, single-cycle permutation in $\Phi_0(n + 1)$. The following lemma resolves this question.

LEMMA 1. *Let ϕ_1 and ϕ_2 be any two permutations in $\Phi(n + 1)$ for which $\phi_1(n + 1) = \phi_2(n + 1) = n + 1$. If the procedure described above yields the same ψ when applied to both ϕ_1 and ϕ_2, then $\phi_1 = \phi_2$.*

To prove this, let $q_1 > \cdots > q_N$ and $r_1 > \cdots > r_M$ be the largest members of the N cycles in ϕ_1 and the M cycles in ϕ_2 respectively, in addition to the 1-cycle $(n + 1)$. Since it is assumed that the above procedure applied to both ϕ_1 and ϕ_2 yields ψ, we have $\psi(n + 1) = \phi_1(q_1) = \phi_2(r_1)$, $\psi(q_i) = \phi_1(q_{i+1})$, $i = 1, \cdots, N - 1$, $\psi(r_j) = \phi_2(r_{j+1})$, $j = 1, \cdots, M - 1$, and $\psi(q_N) = \psi(r_M) = n + 1$.

Further, among all integers $1 \leq k \leq n$, other than those among $(q_1, \cdots, q_N, r_1, \cdots, r_M)$, $\psi(k) = \phi_1(k) = \phi_2(k)$. In particular, $\phi_1(k) = \phi_2(k)$ for all $k < \min(q_N, r_M)$. But now $q_N = r_M$, since $\psi(q_N) = \psi(r_M)$ above. Since q_N is the largest member of two cycles all of whose other members are identical, the cycles are identical and $\phi_1(q_N) = \phi_2(q_N)$. Similarly, since we now have $\psi(q_{N-1}) = \phi_1(q_N) =$

$\phi_2(r_M) = \psi(r_{M-1})$, $q_{N-1} = r_{M-1}$. But $\phi_1(k) = \phi_2(k)$ for all $k < q_{N-1}$, and again q_{N-1} is the largest member of two cycles all of whose other members are identical. Hence, the cycles are identical and $\phi_1(q_{N-1}) = \phi_2(q_{N-1})$. Continuing in this manner then yields the desired result.

Thus, since all members of $\Phi(n) \times \{n + 1\}$ are distinct, the above procedure defines a positive correspondence between $\Phi(n) \times \{n + 1\}$ and the subset of $\Phi_0(n + 1)$ whose members possess a single $n + 1$-cycle. Combining all results above, we conclude that $\Phi^*(n + 1) > \Phi_0(n + 1)$, which by induction now establishes Theorem 5.

Since establishing a positive correspondence between the sets of permutations Φ^* and Φ_0 is sufficient to conclude $M(\phi^*) \geq 1$, it is interesting to consider when such a correspondence does not occur and if the above inequality is violated. Suppose then $\phi \in \Phi(n)$ is arbitrary and let $\Phi^*(\phi)$ and $\Phi_0(\phi)$ be the sets of permutations corresponding to Φ^* and Φ_0 associated with $\phi^* = (1, \cdots, n)$. We are interested in the number of permutations in each set whose ith coordinate equals $\phi(j)$, where i and j are fixed and $i \neq j$. The number of such permutations in $\Phi_0(\phi)$ is easily seen by symmetry to be

(4.6) $$N_0 = D_n/(n - 1),$$

where D_n, of course, is the size of $\Phi_0(\phi)$. The number of permutations in $\Phi^*(\phi)$ whose ith coordinate is $\phi(i)$, however, is given by $\sum_{k=1}^{n-1} k D_{n-1,k}$, since there are k copies of all permutations which match with ϕ in exactly k places among all $j \neq i$, for $1 \leq k \leq n - 1$. Hence, by symmetry again, the number of permutations in $\Phi^*(\phi)$ whose ith coordinate is $\phi(j)$, is

$$N = (D_n - \sum_{k=1}^{n-1} k D_{n-1,k})/(n - 1).$$

From (4.4) it is easily shown that $N = (D_n - (n - 1)!)/(n - 1)$, which combines with (4.6) to yield $N_0 - N = (n - 2)!$. Essentially then, we have $(n - 2)!$ excess permutations in $\Phi_0(\phi)$, whose ith coordinate has been displaced by the jth coordinate of ϕ. In particular, if $i = n$, and $\phi(j) = n$ for some $j < n$, we can conclude that a positive correspondence between $\Phi^*(\phi)$ and $\Phi_0(\phi)$ is impossible. That is, since there are more permutations in Φ_0 than in Φ^* whose nth coordinate is n, some $\phi_1 \in \Phi_0$ must be made to correspond with some $\psi_2 \in \Phi^*$, where $\psi_1(n) = n$ and $\psi_2(n) < n$. In this case, either $\psi_1 > \psi_2$ or the two are incomparable, and hence $\Phi^*(\phi) \not> \Phi_0(\phi)$. Similarly, if $i = 1$ and $\phi(j) = 1$ for $j > 1$, it is again impossible to have a positive correspondence between $\Phi^*(\phi)$ and $\Phi_0(\phi)$, since there are more permutations in Φ_0 than in Φ^* whose first coordinate is 1. Thus, unless ϕ is such that $\phi(1) = 1$ and $\phi(n) = n$ a positive correspondence between $\Phi^*(\phi)$ and $\Phi_0(\phi)$ cannot hold and the following example illustrates that the condition $M(\phi) \geq 1$ will be violated.

Suppose that $(\mathbf{x}; \mathbf{y})$ is a broken sample from a bivariate normal distribution with zero means, unit variances, and correlation $\rho > 0$, and such that $x_1 < \cdots < x_n$ and $y_1 < \cdots < y_n$. Let a and b be nonzero with $a < b$, and let $\delta > 0$ be given.

Suppose further that $x_n = y_n = b$ and that for all $i = 1, \cdots, n - 1$ and $j = 1, \cdots, n - 1$, $|f(x_i, y_j) - A| \leqq A\delta$, $|f(x_i, y_n) - C| \leqq C\delta$ and $|f(x_n, y_j) - C| \leqq C\delta$ where $A = f(a, a)$ and $C = f(a, b) = f(b, a)$. Denoting $f(b, b)$ by B, we have by definition of N and N_0 that for any ϕ such that $\phi(n) < n$,

$$\sum_\phi \{p(\psi) : \psi \in \Phi^*(\phi)\} - \sum_\phi \{p(\psi) : \psi \in \Phi_0(\phi)\}$$

(4.7)
$$\leqq NBA^{n-1}(1 + \delta)^{n-1} + (D_n - N)C^2 A^{n-2}(1 + \delta)^n$$
$$- N_0 BA^{n-1}(1 - \delta)^{n-1} - (D_n - N_0)C^2 A^{n-2}(1 - \delta)^n$$
$$= A^{n-2}(C^2 - AB)(N_0 - N) + \varepsilon,$$

where the quantity ε approaches zero with δ. But $C^2 - AB \propto \exp[-(a^2 - 2\rho ab + b^2)/(1 - \rho^2)] - \exp[-(a^2 + b^2)/(1 + \rho)]$, which is negative since $\rho > 0$, and a and b are nonzero. Hence, for sufficiently small δ, the right side of (4.7) is negative, and $M(\phi) < 1$.

The point of the above discussion is not necessarily surprising in view of Theorem 2 of [1] and Theorem 2 in this paper; namely, that if ϕ maximizes the expected number of correct pairings, then $\phi(1) = 1$ and $\phi(n) = n$. However, if a broader interpretation is permitted, it is suggested that when sampling from an arbitrary MLR distribution, the matchmaker may do rather poorly (worse than random pairing) *unless* he pairs the extreme x- and y-values, i.e. x_1 with y_1 and x_n with y_n.

REFERENCES

[1] DeGroot, M. H., Feder, P. I. and Goel, P. K. (1971). Matchmaking. *Ann. Math. Statist.* **42** 578–593.
[2] Feller, W. (1957). *An Introduction to Probability Theory and its Applications*, **1**. Wiley, New York.
[3] Hardy, Littlewood, and Pólya (1952). *Inequalities.* Cambridge Univ. Press.
[4] Riordan, J. (1958). *An Introduction to Combinatorial Analysis.* Wiley, New York.

School of Management
Rensselaer Polytechnic Institute
Troy, New York 12181