Indian J. pure appl. Math., 15(9) : 968-977, September 1984

ON MINIMUM INTERDEPENDENCE PRINCIPLE

J. N. KAPUR

Department of Mathematics, Indian Institute of Technology, Kanpur 208016

(*Received* 2 *February* 1983; *after revision* 17 *April* 1984)

Some multivariate distributions including the Dirichlet and allied distributions have been obtained by using the Minimum Interdependence Principle (MIP) and these distributions have been used to bring out the different contexts in which MIP and MEP (Maximum Entropy Principle) can be applied. In the discrete case MIP is applied to transportation and international trade problems.

1. INTRODUCTION

Guiasu (1977, 1979) introduced the Principle of Minimum Interdependence (MIP) based on the measure of interdependence introduced by Watanabe (1969). Later Guiasu *et al.* (1980) used the principle for modelling the random molecular networks and Guiasu *et al.* (1982) gave a general solution for MIP distributions both for the discrete and the continuous cases. In particular, they applied MIP to the case of normal variates both when the only constraints are the covariances between the given random variates and when the marginal moments of the first two orders are introduced in the set of constraints. However in the second case, they took the means and variances of the joint distribution the same as the means and variances of the given normal variates. We consider the more general case when these can be different.

We also obtain some more MIP distributions to bring out the distinction between MIP and MEP distributions and the relationship between MIP and MEP measures of interdependence. Incidentally we characterise Dirichlet and allied distributions as MIP distributions. We also consider the application of discrete version of MIP principle to transportation and international trade problems.

2. TWO SIMPLE MIP DISTRIBUTIONS

(i) Let x and y be rectangular variates over the interval [0, 1] and let $E[\ln(x, y)]$ be prescribed as k, then MIP density function is

$$f(x, y) = A \exp[c \ln(xy)] = A(xy)^c \qquad \ldots(1)$$

where

$$A \int_0^1 \int_0^1 (xy)^c \, dx \, dy = 1, \; A \int_0^1 \int_0^1 (xy)^c \ln (xy) \, dx \, dy = k. \qquad \ldots(2)$$

Now

$$\int_0^1 \int_0^1 (xy)^c \, dx \, dy = \frac{1}{(c+1)^2}, \text{ when } c + 1 > 0. \qquad \ldots(3)$$

Differentiating both sides of (3) with respect to c, we get

$$\int_0^1 \int_0^1 (xy)^c \ln (xy) \, dx \, dy = -\frac{2}{(c+1)^3}. \qquad \ldots 4)$$

From (2), (3) and (4)

$$A (c+1)^2, -\frac{2}{c+1} = k \qquad \ldots(5)$$

and

$$(x, y) = \frac{4}{k^2} \left(xy \right)^{\frac{2}{k} - 1}; \; 0 \leqslant x, y \leqslant 1 . \qquad \ldots(6)$$

This MIP distribution exists for all $k < 0$.

The marginal density functions are :

$$\phi_1 (x) = -\frac{2}{k} x^{\frac{2}{k} - 1}, \; 0 \leqslant x \leqslant 1; \; \phi_2 (y) = -\frac{2}{k} y^{\frac{2}{k} - 1}, \; 0 \leqslant y \leqslant 1. \qquad \ldots(7)$$

These are in general different from the rectangular density functions with which we started. If however $k = -2$, the marginal density functions will be the same as the original density functions.

Since for MEP, we insist on marginal density functions being the same as the original density functions, the MEP distribution exists only if $k = -2$ and the constraint $E (\ln (xy)) = k$ becomes ineffective in this case.

(ii) In the above example, let $E (xy)$ be prescribed, so that we get

$$f(x, y) = B \exp(pxy),\ B \int_0^1 \int_0^1 \exp(pxy)\, dx\, dy = 1$$

$$B \int_0^1 \int_0^1 xy \exp(pxy)\, dx\, dy = k \qquad \ldots(8)$$

$$\frac{\int_0^1 \int_0^1 xy \exp(pxy)\, dx\, dy}{\int_0^1 \int_0^1 \exp(pxy)\, dx\, dy} = k. \qquad \ldots(9)$$

If $k < 1$, we can find p from (9) and then determine B and $f(x,y)$ from (8). The marginal density functions are given by

$$\phi_1(x) = B \frac{e^{px} - 1}{px},\ \phi_2(y) = B \frac{e^{py} - 1}{py}. \qquad \ldots(10)$$

These will coincide with the original density functions if $p = 0$, $B = 1$, $k = 1/4$. Again we find that the MIP distribution exists for all $k < 1$, but MEP distribution exists only when $k = 1/4$ and then the additional constraint again becomes ineffective.

3. The Case of Normal Distributions

Let x and y be $N(m_1, s_1^2)$, $N(m_2, s_2^2)$ variates and for the MIP joint density function, let

$$E(x) = \mu_1,\ E(y) = \mu_2,\ E(x^2) = \sigma_1^2 + \mu_1^2,\ E(y^2) = \sigma_2^2 + \mu_2^2,$$

$$E(xy) = \rho\sigma_1\sigma_2 + \mu_1\mu_2 \qquad \ldots(11)$$

then the MIP density function is

$$f(x, y) = A \exp\left[-\frac{1}{2}\frac{(x - m_1)^2}{s_1^2} - \frac{1}{2}\frac{(y - m_2)^2}{s_1^2}\right] \times \exp[-c_1 x^2 - c_2 y^2 - c_3 xy - c_4 x - c_4 y]. \qquad \ldots(12)$$

Determining the constants by using (11) and the fact that $f(x, y)$ is a density function

$$f(x, y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1 - \rho^2}} \exp\left[-\frac{1}{2(1 - \rho^2)}\right] \times \left[\frac{(x - \mu_1)^2}{\sigma_1^2} - \frac{2\rho\,(x - \mu_1)(y - \mu_2)}{\sigma_1\sigma_2} + \frac{(y - \mu_2)^2}{\sigma_2^2}\right] \qquad \ldots(13)$$

which is a bivariate normal distribution whose parameters are independent of m_1, m_2, s_1^2 s_2^2.

Thus the MIP distribution subject to constraints (11) is same as MEP distribution subject to the same constraints.

If we however insist that the marginal density functions of (13) should be the same as the original density functions, then

$$\mu_1 = m_1,\ \mu_2 = m_2,\ s_1^2 = \sigma_1^2,\ s_2^2 = \sigma_2^2. \qquad \ldots(14)$$

Thus MEP distribution satisfying the marginal density constraints exists only in a special case but the MIP distribution always exists though it is independent of the means and variances of the original distributions.

The MIP measure of interdependence is given by

$$[I]_{\text{MIP}} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x, y) \ln \frac{f(x, y)}{\psi_1(x)\,\psi_2(y)}\, dx\, dy \qquad \ldots(15)$$

where $\psi_1(x)$, $\psi_2(y)$ are the density functions of the original normal variates. Simplifying (15), we get

$$[I]_{\text{MIP}} = \ln\left(\frac{s_1}{\sigma_1^2}\right) + \frac{1}{2}\left(\frac{\sigma_1^2}{s_1^2} - 1\right) + \ln\left(\frac{s_2}{\sigma_2}\right) + \frac{1}{2}\left(\frac{\sigma_2^2}{s_2^2} - 1\right)$$

$$- \ln\sqrt{1 - \rho^2} + \frac{1}{2}\frac{(\mu_1 - m_1)^2}{s_1^2} + \frac{1}{2}\frac{(\mu_2 - m_2)^2}{s_2^2} \qquad \ldots(16)$$

$$[I]_{\text{MEP}} = -\ln\sqrt{1 - \rho^2}. \qquad \ldots(17)$$

Now let

$$y = -\ln x + \frac{1}{2}\left(x^2 - 1\right) \qquad \ldots(18)$$

so that

$$\frac{dy}{dx} = -\frac{1}{x} + x,\ \frac{d^2y}{dx^2} = 1 + \frac{1}{x^2}. \qquad \ldots(19)$$

Thus y is a convex function of x which approaches ∞ as $x \to 0$, is equal to 0 at $x = 1$ and approaches ∞ as $x \to \infty$ (Fig. 1) so that $y \geqslant 0$ when $x \geqslant 0$ and $y = 0$ iff $x = y$. In particular $\ln\frac{s_i}{\sigma_i} + \frac{1}{2}\left(\frac{\sigma_i^2}{s_i^2} - 1\right) \geqslant 0$ and equality sign holds iff $s_i = \sigma_i$, $(i = 1, 2)$, $\qquad \ldots(20)$

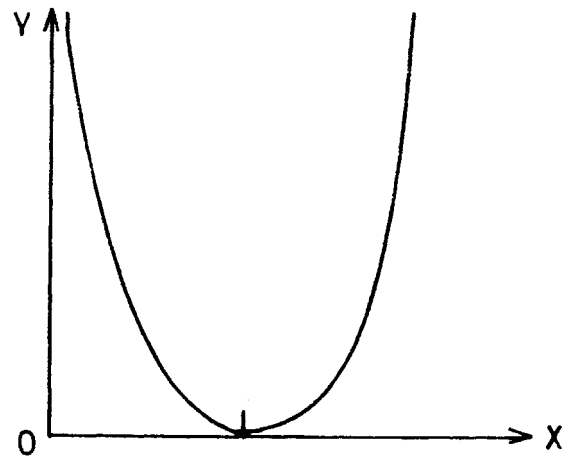

FIG. 1.

From (16), (19) and (20)

$$[I]_{\text{MIP}} \geqslant [I]_{\text{MEP}} \qquad \ldots(21)$$

and equality sign holds iff (14) is satisfied.

4. INEQUALITY BETWEEN I_{MIP} AND I_{MEP}

Now

$$[I]_{\text{MIP}} = \int \ldots \int f(x_1, x_2, \ldots, x_n) \ln \frac{f(x_1, x_2, \ldots, x_n)}{\psi_1(x)\,\psi_2(x_2) \ldots \psi_n(x_n)}\, dx_1, dx_2 \ldots dx_n \qquad \ldots(22)$$

$$[I]_{\text{MEP}} = \int \ldots \int f(x_1, x_2, \ldots, x_n) \ln \frac{f(x_1, x_2, \ldots, x_n)}{\phi_1(x_1)\,\phi_2(x_2) \ldots \phi_n(x_n)}\, dx_1\, dx_2 \ldots dx_n. \qquad \ldots(23)$$

Here $\psi_i(x_i)$ and $\phi_i(x_i)$ are the original and the marginal density functions

$$[I]_{\text{MIP}} - [I]_{\text{MEP}}$$

$$= \int \ldots \int f(x_1, x_2 \ldots, x_n) \ln \frac{\phi_1(x_1)\,\phi_2(x_2) \ldots \phi_n(x_n)}{\psi_1(x_1)\,\psi_2(x_2) \ldots \psi_n(x_n)}\, dx_1\, dx_2 \ldots dx_n$$

$$= \sum_{i=1}^{n} I(\phi_i : \psi_i). \qquad \ldots(24)$$

Thus MIP measure of interdependence exceeds MEP measure of interdependence by the sum of directed divergences of marginal density functions from the original density functions and the two measures are the same iff the marginal density functions are the same as the original density functions.

This shows that any multivariate distribution is closer to the product density of its marginal distributions than to the product density of any other distributions of individual variates.

5. The Dirichlet and Associated Distributions

Let $x_1, x_2, \ldots, x_{n-1}$ have power density functions $m_1 x_1^{m_1-1}, m_2 x_2^{m_2-1}, \ldots,$ $m_{n-1} x_{n-1}^{m_{n-1}-1}$ over the interval [0, 1] and let $E[\ln (1 - x_1 - x_2 \ldots x_{n-1})]$ be prescribed. It is easily seen that the MIP distribution is given by

$$\frac{\Gamma(m_1 + m_2 + \ldots + m_n)}{\Gamma(m_1)\,\Gamma(m_2) \ldots \Gamma(m_n)} x_1^{m_1-1} x_2^{m_2-1} \ldots x_n^{m_n-1} (1 - x_1 - x_2 \ldots x_{n-1})^{m_n-1} \qquad \ldots(25)$$

which is the well-known Dirichlet distribution defined ever

$$0 \leqslant x_i \leqslant 1; x_1 + x_2 + \ldots + x_{n-1} \leqslant 1; i = 1,2,\ldots, n-1. \qquad \ldots(26)$$

The marginal density functions are beta distributions with parameters

$$(m_1, m_2 + \ldots + m_n), (m_2, m_1 + m_3 + \ldots + m_n), (m_3, m_1 + m_2 + m_4 + \ldots + m_n), \ldots,$$
$$(m_{n-1}, m_1 + m_2 + \ldots + m_{n-2}\, m_n)$$

and these are different from the original density functions unless

$$m_1 = m_2 = \ldots = m_{n-1} = \frac{1 - m_n}{n-2}. \qquad \ldots(27)$$

Now

$$I(\phi_1; \psi_1) = \int_0^1 \frac{1}{B(m_1, m_2 + m_3 + \ldots + m_n)} \left[x_1^{m_1-1} (1-x_1)^{m_2+m_3+\cdots+m_n-1} \cdot \right.$$
$$\left. \times \ln \frac{B(m_1, 1)\,(1-x_1)^{m_1+m_2+\cdots m_n-1}}{B(m_1, m_2 + \ldots + m_n)} \right] dx_1 . \qquad \ldots(28)$$

Now using

$$\int_0^1 x^{m-1}(1-x)^{n-1} \ln(1-x)\, dx = \frac{\partial}{\partial n} B(m, n) \qquad \ldots(29)$$

to simplify (28), we get

$$I(\phi_1; \psi_1) = \ln \frac{\Gamma(m_1)\,\Gamma(1)\,\Gamma(m_1 + m_2 + \ldots + m_n)}{\Gamma(m_1+1)\,\Gamma(m_1)\,\Gamma(m_2 + m_3 + \ldots + m_n)} + (m_2 + m_3 + \ldots + m_n - 1)\left[\frac{\Gamma'(m_1)}{\Gamma(m_1)} + \frac{\Gamma'(m_2+\ldots+m_n)}{\Gamma(m_2+\ldots+m_n)}\right] \qquad \ldots(30)$$

$$\begin{aligned}\sum_{i=1}^{n} I(\phi_i; \psi_i) &= -\sum_{i=1}^{n-1} \ln \Gamma(m_i + 1) + (n-1) \ln \Gamma(m_1 + m_2 + \ldots + m_n) \\ &\quad - \sum_{i=1}^{n} \ln \Gamma(m_1 + m_2 + \ldots + m_n - m_i) \\ &\quad + \sum_{i=1}^{n-1} (m_1 + m_2 + \ldots + m_n - m_i - 1)\,[\psi(m_i) \\ &\qquad + \psi(m_1 + m_2 + \ldots + m_n - m_i] \\ &= -\sum_{i=1}^{n-1} \ln \Gamma(m_i+1) + (n-1) \ln \Gamma(M) - \sum_{i=1}^{n-1} \ln \Gamma(M - m_i) \\ &\quad + \sum_{i=1}^{n-1} (M - m_i - 1)\,[\psi(m_i) + \psi(M - m_i)] \qquad \ldots(31)\end{aligned}$$

where

$$M = \sum_{i=1}^{n} m_i. \qquad \ldots(32)$$

If (27) is satisfied, then $m_1 = m_2 = \ldots = m_{n-1} = m$, $m_n = 1 - (n-2)m$

$$M = m(n-1) + 1 - (n-2)m = 1 + m \qquad \ldots(33)$$

$$\begin{aligned}\sum_{i=1}^{n} I(\phi_i; \psi_i) &= -(n-1) \ln \Gamma(m+1) \\ &\quad + (n-1) \ln \Gamma(m) - (n-1) \ln \Gamma(1) = 0. \qquad \ldots(34)\end{aligned}$$

This is expected.

If the range is $[0, \infty)$, then the power density function $A_1 x_1^{m_1-1}$, $A_2 x_2^{m_2-1}$, ..., $A_{n-1} x_{n-1}^{m_n-1}$ are 'improper'. However if $E \ln (1 + x_1 + x_2 + \ldots + x_{n-1})$ is prescribed, then MIP gives the proper MIP density function

$$\frac{\Gamma (m_1 + m_2 + \ldots + m_n) x_1^{m_1-1} x_2^{m_2-1} \ldots x_{n-1}^{m_{n-1}-1}}{\Gamma (m_1) \Gamma (m_2) \ldots \Gamma (m_n) (1 + x_1 + x_2 + \ldots + x_{n-1})^{m_n-1}} \qquad \ldots(35)$$

which is the inverted Dirichlet distributions. By substituting $x_i = e^{-x_i} \cdot i$ in (35) we can get the multivariate logistic distribution.

6. The Transporation and International Trade Problems

Wilson (1970) solved the problem of estimating the trip proportions $p_{ij} = T_{ij}/T$, where T_{ij} is the number of trips from the ith origin to the jth destination and $T = \sum_{j=1}^{n} \sum_{i=1}^{m} T_{ij}$, by maximising the entropy $- \sum_{j=1}^{n} \sum_{i=1}^{m} p_{ij} \ln p_{ij}$ subject to

$$\sum_{j=1}^{n} p_{ij} = p_{i\cdot} \times \sum_{i=1}^{m} p_{ij} = p_{\cdot j} \; (i = 1, 2, \ldots, m, j = 1, 2, \ldots, n) \qquad \ldots(36)$$

$$\sum_{j=1}^{n} \sum_{i=1}^{m} p_{ij} c_{ij} = c \qquad \ldots(37)$$

to get

$$p_{ij} = A_i B_j p_{i\cdot} p_{\cdot j} \exp (-\nu c_{ij}) \qquad \ldots(38)$$

where A_i's, B_j's and ν are determined by using (36) and (37). If the constraints (36) are not insisted upon, we can use MIP to find the joint density function closest to the given product density function $p'_{i\cdot}$, $p'_{\cdot j}$ subject to (37) to get

$$p_{ij} = \frac{p'_{i\cdot} p'_{\cdot j} \exp (-\nu' c_{ij})}{\sum_{i=1}^{n} \sum_{i=1}^{m} p'_{i\cdot} p'_{\cdot i} \exp (-\nu' c_{ij})} \qquad \ldots(39)$$

where ν' is obtained by using (37). Thus MIP distribution will not necessarily satisfy the constraints (36).

This will be useful when we know the proportion of trips originating at each origin and also the proportions terminating at each distination. However the current proportions need not be known.

Alternatively we use MDI principle to minimize

$$\sum_{j=1}^{n} \sum_{i=1}^{m} p_{ij} \ln \frac{p_{ij}}{p'_{i.}\, p'_{.j}} \qquad \ldots(40)$$

subject to (36) and (37) to get

$$p_{ij} = c_i D_j p'_{i.} p'_{.j} \exp [-\nu'' c_{ij}]. \qquad \ldots(41)$$

If we use the constraints (36) and (37) to determine the constaints, we shall get the same solution as (38) and the final solution will be independent of $p'_{i.}$ and $p'_{.j}$.

Similar problems arise in international trade (Theil 1967, Kapur 1982, 1984) for estimating the trade between the ith jth countries when imports and exports in earlier years are known.

7. Concluding Remarks

The MIP can be used to generate multivariate density functions of the form

$$A\psi_1(x_1)\,\psi_2(x_2)\ldots\psi_n(x_n) \exp\left[-\sum_{r=1}^{m} \nu_r g_r(x_1, x_2, \ldots, x_n)\right]. \qquad \ldots(42)$$

This is the joint density function which is nearest to the product density function $\psi_1(x_1)\,\psi_2(x_2)\ldots\psi_n(x_n)$ and for which the constraints.

$$E[g_r(x_1, x_2, \ldots, x_n)] = a_r,\ r = 1, 2, \ldots, m \qquad \ldots(43)$$

are imposed. For this distribution, the original density functions $\psi_1(x_1)$, $\psi_2(x_2)$,..., $\psi_n(x_n)$ may not be the marginal density functions $\phi_1(x_1)$, $\phi_2(x_2)$,...,$\phi_n(x_n)$. The directed divergence of the density (42) from the product density $\psi_1(x_1)\,\psi_2(x_2)$..., $\psi_n(x_n)$ exceeds its directed divergence from the product density $\phi_1(x_1)\,\phi_2(x_2)\ldots\phi_n(x_n)$ by the sum of directed divergenes of each ϕ_i from ψ_i.

We can obtain distributions of the from (42) in the following different ways :

(i) start with the product density $\psi_1(x_1)\,\psi_2(x_2)\ldots\psi_n(x_n)$, use the constraints (43) and apply MIP.

(ii) use constraints (43) and such other constraints as make $f_1(x_1), f_2(x_2),\ldots,f_n(x_n)$ maximum entropy distributions and apply MEP.

(iii) start with density function $\psi_1(x_1)\,\psi_2(x_2)\ldots\psi_n(x_n)$ and use all the constraints in (ii) and use MIP to get a density function which is independent of $\psi_1(x_1), \psi_2(x_2),\ldots,\psi_n(x_n)$.

REFERENCES

Guiasu, S. (1977). Information Theory with Applications. McGraw-Hill Book Co., Inc., New York.

——— (1979). On entropy measure of connection and interdependence between the subsystems of a given large system. In Proceedings of Symposium on Mathematical Methods for the Analysis of Large-Scale Systems. Academic Publishing House, Czech Academy of Sciences, Prague.

Guiasu, S., Martin, L, and Reischer, C. (1980). On the structure of random modular network. *J. Camp. Inf. Syst. Sci.* 5, 205–14.

Guiasu, S., Leblanc, R., and Reischer, C. (1982). Principle of minimum interdependence. *J. Inf. Opt. Sci.* 3, (2), 149–72.

Kapur, J. N. (1982). Lecture Notes on Large Scale Systems. Waterloo University, Canada.

——— (1984). Maximum-Entropy Models in Science and Engineering, South Asia Publishers, New Delhi.

Theil, H., (1967). Economics and Information Theory. North Holland, Amestardam.

Wilson, A. G. (1970). Entropy in Urban and Regional Modelling. Pion, London.

Watanabe, S. (1969). Knowing and Guessing. John Wiley, New York.