Journal of Sciences, Islamic Republic of Iran 16(4): 339-345 (2005) http://jsciences.ut.ac.ir
University of Tehran, ISSN 1016-1104

Determination of Maximum Bayesian Entropy Probability Distribution

S. Mansoury,[1] E. Pasha,[2] and M. Mohammadzadeh[1,*]

[1] *Department of Statistics, Faculty of Basic Sciences, Tarbiat Modarres University, P.O. Box 14115-175, Tehran, Islamic Republic of Iran*
[2] *Department of Statistics, Faculty of Basic Sciences, Tarbiat Moalem University, Tehran, Islamic Republic of Iran*

Abstract

In this paper, we consider the determination methods of maximum entropy multivariate distributions with given prior under the constraints, that the marginal distributions or the marginals and covariance matrix are prescribed. Next, some numerical solutions are considered for the cases of unavailable closed form of solutions. Finally, these methods are illustrated via some numerical examples.

Keywords: Bayesian entropy; Maximum entropy principle; Splines; Curve fitting

1. Introduction

Bayesian entropy of a pdf (pmf) f with given prior pdf (pmf) α is defined as

$$B(f) = -E_f\left(\ln\frac{f(X)}{\alpha(X)}\right) \tag{1}$$

The maximum Bayesian entropy probability distributions (MEPD) were studied by a number of researchers such as Jaynes [9], Ingarden and Kassakowski [7], Jain and Consul [8] Guiasu [4], Consul and Shenton [2], Consul and Jain [1], Hobson and Cheng [6], Tribus and Rossi [14], Georgescu [3], Guiasu [5], Kapur [10]. Mansoury, *et al.* [12] gave a general form of the maximum entropy bivariate probability distributions (MEBPD) via Shannon's measure of entropy, when the marginal probability density functions (pdf) are prescribed and it was extended for the maximum entropy multivariate probability distributions (MEMPD). In this paper determination of MEMPD with given marginal distributions or marginals and variance and covariance matrix, via Bayesian entropy is considered in section 2. Next two numerical methods by using power extension and smooth curve fitting are given in section 3. Results and discussions are given in the final section.

2. Determination of MEMPD

In this section, to determine the MEMPD with given marginals, covariance matrix and prior function, the following lemma and theorems are presented.

Lemma 1. Let L^2 denotes the collection of all measurable and squared integrable functions. If h_1 and h_2 are two pdfs in L^2, then $h_1 = h_2$; a.e. if and only if, for any arbitrary function K,

$$\int_R K(x)h_1(x)dx = \int_R K(x)h_2(x)dx \tag{2}$$

Proof. When $h_1 = h_2$, a.e. then for any arbitrary

[*] *E-mail: mohsen_m@modares.ac.ir*

function K, Equation (2) is obvious. For the converse, from (2)

$$\int_R K(x)(h_1(x)-h_2(x))dx = 0 .$$

Since K is arbitrary, for $K(x)=h_1(x)-h_2(x)$, we have

$$\int_R (h_1(x)-h_2(x))^2 dx = 0 ,$$

Therefore $h_1 = h_2$; a.e.

Theorem 1. Let $g_1, \cdots, g_n \in L^2$, be prescribed marginal pdfs of multivariate distribution of $\mathbf{X}=(\mathrm{X}_1,\cdots,\mathrm{X}_n)$ with given prior $\alpha(x)$, $\mathbf{x}=(x_1,\cdots,x_n)\in E \subseteq R^n$. The MEMPD can be uniquely obtained by

$$f_M(\mathbf{x}) = \alpha(\mathbf{x})\prod_{i=1}^{n} f_i(x_i), \quad \mathbf{x}\in \mathrm{E}, \tag{3}$$

in which, the functions $f_1, \cdots, f_n$ are found by solving the following system of functional equations

$$\int_{R^{n-1}} \alpha(\mathbf{x}) \prod_{i=1}^{n} f_i(x_i) d\mathbf{x}_{-i} = g_i(x_i); \; i=1,\cdots,n, \tag{4}$$

where $d\mathbf{x}_{-i}$ denotes $dx_1\cdots dx_{i-1}dx_{i+1}\cdots dx_n$.

Proof. Since the marginal pdfs of $f(\mathbf{x})$ are prescribed, for any function $K_i(X_i)$,

$$\mu_i = E(K_i(X_i)); \; i=1,\cdots,n \tag{5}$$

are known. By Lemma 1, prescribed $g_1,\cdots,g_n$ are equivalent to prescribe $EK_i(X_i)$ for any function $K_i(X_i)$. Now MEMPD, can be found by maximization of (1) with constraints (5). By Euler-Lagrange method (Kapur, 1989), the Lagrangian is given by

$$L = \int_E (-f)\ln\frac{f}{\alpha}d\mathbf{x} + \sum_{i=1}^{n}\lambda_i\left(\int_E K_i(x_i)fd\mathbf{x}-\mu_i\right)$$

$$= \int_E \{-f \ln\frac{f}{\alpha} + \sum_{i=1}^{n}\lambda_i K_i(x_i)f\}d\mathbf{x} - \sum_{i=1}^{n}\lambda_i\mu_i .$$

Since $(-f)\ln\frac{f}{\alpha}+\sum_{i=1}^{n}\lambda_i K_i(x_i)f$ is a concave and continuous function of f, the unique extermal solution maximizes the entropy. Therefore

$$\frac{\partial}{\partial f}\{(-f)\ln\frac{f}{\alpha}+\sum_{i=1}^{n}\lambda_i K_i(x_i)f\} = 0 ,$$

and MEMPD is given by

$$f_M(\mathbf{x}) = \alpha(\mathbf{x})\exp\{\lambda_1 K_1(x_1)\}\cdots\exp\{\lambda_n K_n(x_n)\} .$$

It is clear that f_M is the product of $n+1$ separate functions $\alpha(x_1,\cdots,x_n)$ and $f_1(x_1), \cdots, f_n(x_n)$, *i.e.*

$$f_M(\mathbf{x}) = \alpha(\mathbf{x})\, f_1(x_1)\cdots f_n(x_n) \qquad \mathbf{x}\in E .$$

Since f_M has the marginals $g_1, \cdots, g_n$, the functions $f_1, \cdots, f_n$ can be obtained by solving the system of Equations (4).

Theorem 2. Let $g_1, \cdots, g_n \in L^2$ are marginal pdfs of multivariate distribution of $\mathbf{X}$ with prescribed covariance matrix Σ and prior function $\alpha(\mathbf{x})$, with $\mathbf{x}\in E \subseteq R^n$, then the MEMPD is uniquely given by

$$f_M(\mathbf{x}) = \alpha(\mathbf{x})f_1(x_1)\cdots f_n(x_n)e^{(\mathbf{x}-\mu)'\Lambda(\mathbf{x}-\mu)}, \; \mathbf{x}\in E ,$$

Where $\mu = E\mathbf{X}$, and $f_1, \cdots, f_n$ and Λ can be obtained by solving the following system of equations.

Proof. Since the marginal pdfs of $f(\mathbf{x})$ are prescribed, for any function $K_i(X_i)$

$$\mu_{K_i} = E(K_i(X_i)); \quad i=1,\cdots,n \tag{7}$$

are known. Prescribed $g_1,\cdots,g_n$ and covariance matrix are equivalent to prescribed $EK_i(X_i)$ and

$$\sigma_{i,j} = E(X_i-\mu_i)(X_j-\mu_j) \quad i,j=1,\cdots,n \tag{8}$$

where $\mu_i = EX_i$, $i=1,\cdots,n$. Now MEBPD can be fined by maximization of (1) with constraints (7) and (8). The Lagrangian is given by

$$L = \int_E (-f)\ln\frac{f}{\alpha}d\mathbf{x} + \sum_{i=1}^{n}\lambda_i\left(\int_E K_i(x_i)fd\mathbf{x}-\mu_{K_i}\right)$$

$$+\sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_{i,j}\left(\int_E (x_i-\mu_i)(x_j-\mu_j)fd\mathbf{x}-\sigma_{i,j}\right)$$

$$-\sum_{i=1}^{n}\lambda_i\mu_{K_i} - \sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_{i,j}\sigma_{i,j} .$$

Since

$$(-f)\ln\frac{f}{\alpha}+\sum_{i=1}^{n}\lambda_i K_i(x_i)f$$
$$+\sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_{i,j}(x_i-\mu_i)(x_j-\mu_j)f$$

is a concave and continuous function of *f*, the unique extermal solution maximizes entropy. Therefore

$$\frac{\partial}{\partial f}\{-f\ \ln\frac{f}{\alpha}+\sum_{i=1}^{n}\lambda_i K_i(x_i)f\}$$
$$+\sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_{i,j}(x_i-\mu_i)(x_j-\mu_j)f=0,$$

and MEMPD is given by

$$f_M(\mathbf{x})=\alpha(\mathbf{x})\exp\{\lambda_1 K_1(x_1)\}\cdots\exp\{\lambda_n K_n(x_n)\}$$
$$.\exp\{\sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_{i,j}(x_i-\mu_i)(x_j-\mu_j)\}$$
$$=\alpha(\mathbf{x})f_1(x_1)\cdots f_n(x_n)e^{(\mathbf{x}-\mu)'\Lambda(\mathbf{x}-\mu)},$$

where $\Lambda=[\lambda_{i,j}]_{n,n}$. Since f_M has the marginals $g_1,\cdots,g_n$ and covariance matrix $\sum$, the functions $f_1,\cdots,f_n$ are obtained by solving the system of Equations (6).

Theorems 1 and 2 can easily be extended for discrete multivariate random variables by replacing integrals with summations.

3. Numerical Methods

Since the integral and functional Equations (4) and (6) can not be always solved analytically, we consider two numerical methods to obtain the MEMPD in special case. Suppose we want to determine the MEBPD over the region

$$E=\{(x,y)\in R^2 \mid\ a\le x\le b, C_1(x)\le y\le C_2(x)\}$$
$$=\{(x,y)\in R^2 \mid\ c\le y\le d, D_1(y)\le x\le D_2(y)\},$$

(Figure 1) with given marginal pdfs *g* and *h* on (a,b) and (c,d) respectively, and uniform prior distribution over *E*.

By Theorem 1 the general form of MEBPD is given by

$$f_M(x,y)=f_1(x)f_2(y),\quad (x,y)\in E. \tag{9}$$

Here two numerical methods for approximation of f_1 and f_2 are considered, that can easily be extended to approximate MEMPD.

Method 1. Suppose $f_1(x)$ and $f_2(y)$ can be expanded in power series

$$f_1(x)=\sum_{n=0}^{\infty}A_n x^n,\ f_2(y)=\sum_{n=0}^{\infty}B_n y^n.$$

The marginal pdfs of f_M can be written a

$$g(x)=(\sum_{i=0}^{\infty}A_i x^i)\sum_{j=0}^{\infty}\frac{B_j}{j+1}(C_2^{j+1}(x)-C_1^{j+1}(x)), \tag{10}$$

and

$$h(y)=(\sum_{j=0}^{\infty}B_j y^j)\sum_{i=0}^{\infty}\frac{A_i}{i+1}(D_2^{i+1}(y)-D_1^{i+1}(y). \tag{11}$$

Now the coefficients $A_1,A_2,\cdots$ and $B_1,B_2,\cdots$ can be found by joining the Equations (10) and (11) in the following steps.

Step 1. Let $\sum_{n=0}^{\infty}G_n x^n$, $\sum_{n=0}^{\infty}H_n y^n$, $\sum_{n=0}^{\infty}C_n^{k,j}x^n$ and $\sum_{n=0}^{\infty}D_n^{k,j}y^n$ for $k=1,\ 2$ and $j=1,\ 2,\ \cdots$ uniformly converge to $g(x)$, $h(y)$, $C_k^j(x)$ and $D_k^j(y)$ respectively, where $G_n=\frac{1}{n!}\frac{d^n}{dx^n}g(x)\big|_{x=0}$, $H_n=\frac{1}{n!}\frac{d^n}{dy^n}h(y)\big|_{y=0}$, $C_n^{k,j}=\frac{1}{n!}\frac{d^n}{dx^n}C_k^j(x)\big|_{x=0}$, and $D_n^{k,j}=\frac{1}{n!}\frac{d^n}{dy^n}D_k^j(y)\big|_{y=0}$.

Step 2. By taking derivations of Equation (10) and putting $x=0$, for each $k=0,1,\cdots,$ we have

$$g^{(k)}(0)=\sum_{i=0}^{k}\left\{\binom{k}{i}\frac{d^{k-i}}{dx^{k-i}}(\sum_{j=0}^{\infty}A_j x^j)\frac{d^i}{dx^i}\right.$$
$$\left.\left[\sum_{j=0}^{\infty}\frac{B_j}{j+1}(C_2^{j+1}(x)-C_1^{j+1}(x))\right]\right\}\Big|_{x=0},$$

so that

$$G_k = \sum_{i=0}^{k} A_{k-i} \sum_{j=0}^{\infty} \frac{B_j}{j+1}(C_i^{2,j+1} - C_i^{1,j+1}),$$

$$k = 0,1,\cdots, \quad (12)$$

and

$$H_k = \sum_{i=0}^{k} B_{k-i} \sum_{j=0}^{\infty} \frac{A_j}{j+1}(D_i^{2,j+1} - D_i^{2,j+1}),$$

$$k = 0,1,\cdots. \quad (13)$$

Step 3. Now f_1 and f_2 can be approximated by Taylor series with $(n+1)$ terms. The coefficients $A_0, A_1, \cdots, A_n$ and $B_0, B_1, \cdots, B_n$ are obtained from solving the system of Equations (12) and (13) by Newton's method.

Step 4. The approximate MEBPD is now given by

$$\hat{f}_M(x,y) = \hat{f}_1(x)\hat{f}_2(y), \quad (x,y) \in E. \quad (14)$$

Example 1. Let $E = \{(x,y) \in R^2 \mid \ 0 < x < y < 1\}$, $g(x) = 4x(1-x^2)$, $0 < x < 1$ and $h(y) = 4y^3$, $0 < y < 1$. By Step 1 we have

$$G_n = \frac{1}{n!}\frac{d^n}{dx^n}(g(x))\Big|_{x=0}, \ H_n = \frac{1}{n!}\frac{d^n}{dy^n}(h(y))\Big|_{y=0},$$

$$C_n^{k,j} = \frac{1}{n!}\frac{d^n}{dx^n}(C_k^j(x))\Big|_{x=0},$$

$$C_1(x) = x, \ C_2(x) = 1, \ D_1(y) = 0, \ D_2(y) = y$$

and

$$f_1(x) = \sum_{n=0}^{\infty} A_n x^n, \ f_2(y) = \sum_{n=1}^{\infty} B_n y^n.$$

Now by Step 2 the coefficients of expansion for the marginals g and h respectively are given by

$$G_0 = 0, \ G_1 = 4, \ G_2 = 0, \ G_3 = -4, \ G_n = 0, \ n \geq 4$$

$$H_3 = 4, \ H_n = 0, \ n \neq 3$$

Also we have

$$C_n^{1,j} = \begin{cases} 1 & \text{n} = \text{j} \\ 0 & \text{n} \neq \text{j} \end{cases}, \ C_n^{2,j} = \begin{cases} 1 & \text{n} = \text{j} \\ 0 & \text{n} \neq \text{j} \end{cases},$$

$$D_n^{2,j} = \begin{cases} 1 & \text{n} = \text{j} \\ 0 & \text{n} \neq \text{j} \end{cases}, \ D_n^{1,j} = 0, \ n \in N.$$

In Step 3, by solving the nonlinear system of Equations (12) and (13), we get

$$A_i B_j = \begin{cases} 8 & (i,j) = (1,1) \\ 0 & (i,j) \neq (1,1) \end{cases}$$

Therefore by Step 4 we have

$$\hat{f}_M(x,y) = 8xy; \quad (x,y) \in E.$$

Method 2. From (9) the marginal pdfs of f_M are

$$g(x) = f_1(x)\int_{C_1(x)}^{C_2(x)} f_2(y)dy; \quad a < x < b, \quad (15)$$

$$h(y) = f_2(y)\int_{D_1(y)}^{D_2(y)} f_1(x)dx; \quad c < y < d. \quad (16)$$

Now we subdivide E into subregions like Figure 1.

$$a = x_0 < x_1 < \cdots < x_{n_1+1} = b,$$

$$c = y_0 < y_1 < \cdots < y_{n_2+1} = d,$$

$$\Delta x_i = x_i - x_{i-1} = \Delta, \quad i = 2,3,\cdots,n_1$$

$$\Delta y_j = y_j - y_{j-1} = \Delta, \quad j = 2,3,\cdots,n_2$$

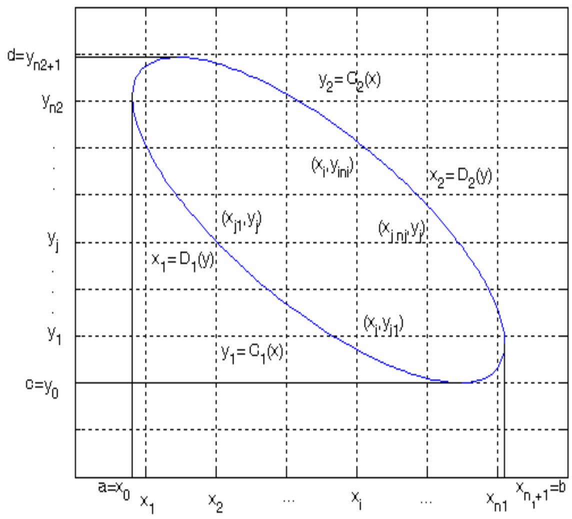

Figure 1. Gridded E in R^2.

From (15) and (16), for each $1 \le i \le n_1$ and $1 \le j \le n_2$, we have

$$g(x_i) = f_1(x_i)\{\frac{f_2(C_1(x_i)) + f_2(y_{i1})}{2}(y_{i1} - C_1(x_i))$$

$$+\sum_{k=2}^{n_i}\frac{f_2(y_{ik}) + f_2(y_{i,k-1})}{2}\ (y_{ik} - y_{i,k-1})$$

$$+\frac{f_2(y_{in_i}) + f_2(C_2(x_i))}{2}(C_2(x_i) - y_{in_i})\ \},$$

$$h(y_j) = f_2(y_j)\{\frac{f_1(D_1(y_j)) + f_1(x_{j1})}{2}(x_{j1} - D_1(y_j))$$

$$+\sum_{k=2}^{n_j}\frac{f_1(x_{jk}) + f_1(x_{j,k-1})}{2}(x_{jk} - x_{j,k-1})$$

$$+\frac{f_1(x_{jn_j}) + f_1(D_2(y_j))}{2}(D_2(y_j) - x_{jn_j})\}.$$

By replacing the following approximations:

$f_2(C_1(x_i)) = f_2(y_{i1}),\ f_2(C_2(x_i)) = f_2(y_{in_i})$
$1 \le i \le n_{1,}$

$f_1(D_1(y_j)) = f_1(x_{j1}), f_1(D_2(y_j)) = f_1(x_{jn_j})$
$1 \le j \le n_{2.},$

we have

$$g(x_i) = f_1(x_i)\{f_2(y_{i1})(y_{i1} - C_1(x_i))$$

$$+\frac{\Delta}{2}[f_2(y_{i1}) + 2(f_2(y_{i2}) + \cdots$$

$$+f_2(y_{i,(n_i-1)})) + f_2(y_{in_i})]$$

$$+f_2(y_{in_i})(C_2(x_i) - y_{in_i})\ \}, \qquad (17)$$

$$h(y_j) = f_2(y_j)\{f_1(x_{i,1})(x_{j,1} - D_1(y_j))$$

$$+\frac{\Delta}{2}[f_1(x_{j1}) + 2(f_1(x_{j2}) + \cdots$$

$$+f_1(x_{j,(n_j-1)})) + f_1(x_{jn_i})]$$

$$+f_1(x_{jn_j})(D_2(y_j) - x_{jn_j})\}. \qquad (18)$$

Since the Equations (17) and (18) are related, we set one of the nonzero and unknowns $f_1(x_i)$ or $f_2(y_j)$ to be equal to 1, then the other unknowns can be found from (17) and (18). Now, we fit two smooth curves $\hat{f}_1$ and $\hat{f}_2$ over intervals (a,b) and (c,d). Then an approximate of MEBPD over E is given by (14). Now three methods for fitting smooth curves to f_1 and f_2 are considered.

Method 2.1. The approximation of a function by cubic spline with not-a matrix -knot end conditions, meaning that it is the unique piecewise cubic polynomial with two continuous derivatives with breaks at all interior data sites except on the leftmost and the rightmost one, is illustrated in Example 2 [13].

Example 2. Let $E = \{(x,y) \in R^2 \,|\, 0 < x + y < 1, 0 < x < 1\}$ and the marginal pdfs of f_M be Beta (2,3). By Method 2, we have

$$g(x) = f_1(x)\int_0^{1-x} f_2(y)dy = \frac{\Gamma(5)}{\Gamma(2)\Gamma(3)}x(1-x)^2, \quad 0 < x < 1$$

$$h(y) = f_2(y)\int_0^{1-y} f_1(x)dx = \frac{\Gamma(5)}{\Gamma(2)\Gamma(3)}y(1-y)^2, \quad 0 < y < 1.$$

Since this problem is symmetric, so $f_1 = f_2 = f$. By letting $x = 0,\ 0.1,\ \cdots,\ 0.9$ and solving the following system of equations

$$f(0)\{f(0)+2[f(0.1)+\cdots+f(0.9)]+f(1)\}=0$$

$$f(0.1)\{f(0)+2[f(0.1)+\cdots+f(0.8)]+f(0.9)\}=19.44$$

$$f(0.2)\{f(0)+2[f(0.1)+\cdots+f(0.7)]+f(0.8)\}=30.72$$

$$\vdots$$

$$f(0.8)\{f(0)+2f(0.1)+f(0.2)\}=7.68$$

$$f(0.9)\{f(0)+f(0.1)\}=2.16$$

we get

$$f(0)=0.0000, f(0.1)=0.4899, f(0.2)=0.9798,$$

$$f(0.3)=1.4697, f(0.4)=1.9596, f(0.5)=2.4495,$$

$$f(0.6)=2.9394, f(0.7)=3.4293, f(0.8)=3.9192,$$

$$f(0.9)=4.4091, f(1)\approx 4.4091.$$

By fitting a cubic spline to this set of points, we have

$$\hat{f}(x)=$$

$$\begin{cases} 0.0075x^3-0.0023x^2+4.8991x & 0\le x\le 0.1\\ 0.0075x^3+4.8989x+0.4899 & 0.1\le x\le 0.2\\ -0.0376x^3+0.0023x^2+4.8991x+0.9798 & 0.2\le x\le 0.3\\ 0.1428x^3-0.0090x^2+4.8985x+1.4697 & 0.3\le x\le 0.4\\ -0.5336x^3+0.0338x^2+4.9009x+1.9596 & 0.4\le x\le 0.5\\ 1.9916x^3-0.1263x^2+4.8917x+2.4495 & 0.5\le x\le 0.6\\ -7.4329x^3+0.4712x^2+4.9262x+2.9394 & 0.6\le x\le 0.7\\ 27.7404x^3-1.7587x^2+4.7974x+3.4293 & 0.7\le x\le 0.8\\ -103.5276x^3+6.5634x^2+5.2779x+3.9192 & 0.8\le x\le 0.9\\ -103.5276x^3-24.4949x^2+3.4848x+4.4091 & 0.9\le x\le 1 \end{cases}$$

Now an approximate of MEBPD over E can be determined by (14) which is shown in Figure (2).

Method 2.2. Suppose we want to fit a Fourier series to $f(x)$; $a\le x\le b$ and $a>0$, by using the following set of data

$$\{(x_i, f(x_i)) \mid i=0,1,\cdots,n_1+1\}. \tag{19}$$

We define

$$g(x)=\begin{cases} f(x) & a\le x\le b\\ 0 & 0\le x\le a \end{cases}$$

and outside of this interval $g(x)=g(x+2l)$ where $2l=a$. The Fourier series corresponding $g(x)$ is given by

$$\frac{A_0}{2}+\sum_{i=1}^{\infty}(A_n Sin\frac{n\pi x}{l}+B_n Cos\frac{n\pi x}{l}).$$

The coefficients can be approximate by

$$\hat{A}_n=\frac{\Delta}{l}\sum_{i=0}^{n_1} f(x_i) Cos\frac{n\pi x_i}{l},$$

$$\hat{B}_n=\frac{\Delta}{l}\sum_{i=0}^{n_1} f(x_i) Sin\frac{n\pi x_i}{l}.$$

Then

$$\hat{f}(x)=\frac{\hat{A}_0}{2}+\sum_{n=1}^{\infty}(\hat{A}_n Sin\frac{n\pi x}{l}+\hat{B}_n Cos\frac{n\pi x}{l}),$$

is an approximation of $f(x)$ over the interval (a,b).

Method 2.3. To expand $f(x); a\le x\le b$ into the Hermite series. We define

$$g(x)=\begin{cases} f(x) & a\le x\le b\\ 0 & 0\le x\le a \end{cases}$$

therefore

$$g(x)=\sum_{n=1}^{\infty} A_n H_n(x)$$

where

$$A_n=\frac{1}{2^n n!\sqrt{\pi}}\int_R e^{-x^2} f(x) H_n(x)dx,$$

$$H_n(x)=(-1)^n e^{x^2}\frac{d^n}{dx^n}e^{-x^2}, \quad n=0, 1,\cdots$$

To fit a Hermite series into the set of data (19), we approximate coefficients by

$$\hat{A}_n=\frac{\Delta}{2^n n!\sqrt{\pi}}\sum_{k=0}^{n_1} e^{-x_k^2} f(x_k) H_n(x_k).$$

Now an approximate of $f(x)$ over the interval (a,b) is given by

$$\hat{f}(x)=\sum_{n=1}^{\infty}\hat{A}_n H_n(x)\ .$$

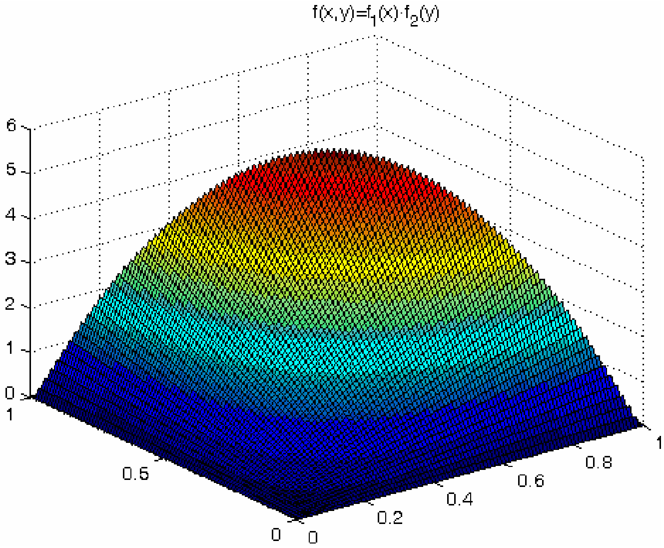

Figure 2. Plot of approximate MEBP distribution.

4. Results and Discussions

The determination methods of MEPD for univariate probability distribution functions are extended for multivariate cases via Bayesian entropy. Two numerical methods for determination of the maximum entropy bivariate probability distributions with unavailable analytical solutions are also given. It should be noted that the solving of the system of equations in the proposed methods for MEMPD via Bayesian entropy is very complicated. Therefore, providing a fast computational method for these cases is required. Since the MEMPD with Shannon's measure of entropy can be obtained by choosing a uniform prior distribution in the Bayesian entropy, it seems that other measures, say Renyi's, Kapur's, Havrada and Charvat's and Burg's measures of entropy can be considered to find MEMPD, that needs further studies.

Acknowledgement

The authors like to thank the Referees of the Journal for their valuable comments and suggestion to refine the article.

References

1. Consul P.C. and Jain G.C. A generalization of the Poisson distribution. *Technometrics,* **15**(4): 791-799 (1973).
2. Consul P.C. and Shenton L.R. Use of lagrange expansion

for generating discrete generalized probability distributions. *SIAM J. App. Math*., **23**(2): 239-248 (1972).
3. Georgescu Roegen N. *The Measures of Information, A Critique, in Modern Trens in Cybernetics and Systems*. Springer Verlag, vol. 2, 187-217 (1975).
4. Guiasu S. Weighted entropy. *Report on Math Physics*, **2**: 165-179 (1971).
5. Guiasu S. *Information Theory with Applications*. McGraw Hill, New York (1977).
6. Hobson A. and Cheng B.K. A comparison of the shannon and kullback information measures. *Journal. Stat. Phys*., **7**: 301-310 (1973).
7. Ingarden R.S. and Kassakowski A. Poisson probability distribution and informations thermodynamics. *Bull. Acad. Polon*. *Polonaise Sci. Math.*, **19**: 86-88 (1971).
8. Jain G.C. and Consul P.C. A generalized negative binomial distribution. *SIAM. J. App. Math.*, **21**: 501-513 (1971).
9. Jaynes E.T. Prior probabilities. *IEEE Trans. S.S.C.*, **4**: 227-248 (1968).
10. Kapur J.N. Bayesian Entropy and its Application to Entropy Maximization on Models. IIT/K Res. Rep. (1983).
11. Kapur J.N. *Maximum Entropy Models in Science and Engineering*. Wiley Eastern Limited, New Delhi, India (1989).
12. Mansoury S., Pasha E., and Mohammadzadeh M. Determination of maximum entropy probability distribution under some constraints, to appear in *Australian Journal of Statistics* (2005).
13. Mathews J.H. and Kurtis D.F. *Numerical Methods Using MATLAB.* Prentice-Hall (1992).
14. Tribus M. and Rossi R. On kullback information measures as a basis for information theory. *Jour. Stat. Phy.*, **9**: 331-338 (1973).