MODE ESTIMATION FOR A BIVARIATE DISTRIBUTION

İ. KARABULUT

Department of Statistics, Ankara University, Ankara, Turkey

ABSTRACT

Given a sequence of independent and identically distributed random vectors $(X_1, Y_1), (X_2, Y_2), (X_3, Y_3), ..., (X_n, Y_n)$ with a unimodal bivariate distribution function $F(x, y)$, a consistent estimator of the mode, (θ_x, θ_y) is proposed by using spacings as defined by J.H. Venter (1967) (Ann. Math. Statist., 38, 1446-1455) for univariate distributions. The method is illustrated for a bivariate distribution.

1. INTRODUCTION

A distribution (d.f.) F of the random variable X is called unimodal at x if the graph of F is convex in $(-\infty, x)$ and concave in (x, ∞). Unimodality requires that there exist a density f which is monotone in $(-\infty, x)$ and (x, ∞), such that constancy are not excluded (see Feller [4], footnote on p. 158). Mode is an important concept, though its estimation has so far received little attention, at least to my knowledge.

The estimation of the mode of a multivariate d.f. may be desirable in some instances. For example, it can be used to find an upper bound estimation of a multivariate hazard function

$$\frac{f(x)}{1 - F(x)}.$$

An another case is to utilize it in defining a skewness measure for an appropriate family of multivariate distributions. It is also possible to enumerate further cases of application where the estimation of the mode of a multivariate distribution is required. Such a case for the heat equation modelled by the Weiner-Lévy proces is given by Chernoff [2] in the univariate setting.

A short review of the studies on the mode estimation for the unimodal univariate distributions will reveal that it is possible to follow different routes. In his seminal paper on the estimation of density functions, Parzen [6] proposes a

method called analogue estimation for the mode of a unimodal univariate d.f. First, a consistent estimation $\hat{f}_n$ of a density function f is obtained by using the definiton of the empirical d.f. Then, the anologue estimation $\hat{\theta}_x$ of the unkown mode θ_x is proposed as

$$\hat{f}_n(\hat{\theta}_x) = \max_{-\infty < x < \infty} \hat{f}_n(x).$$

The consistency of this estimation depends on the uniform continuity of the underlying p.d.f. f(x). The asymptotic normality of this estimator is shown under some mild conditions. In the same line Chernoff [2] develops Parzen [6]'s analog estimators of a continuously differentiable f by using bounded weighting (kernel) function that has bounded first and second derivatives. In the estimation of the mode, he uses the idea that the most of the observations are clustered around a point which is the mode itself θ_x.

On the other hand, Grenander [5] follows different route. Grenander [5] uses the weighted average of order statistics, basically the spacings based on the order statistics $X_{1:n} \leq X_{2:n} \leq X_{3:n} \cdots \leq X_{n:n}$ constructed from a random sample of size n. The weights tend to be large at the values where the density function is large. The mode estimator is given as

$$\hat{\theta}_x = \frac{\sum_{v=1}^{n-k} \frac{1}{2}(X_{v+k:n} + X_{v:n})}{\sum_{v=1}^{n-k}(X_{v+k:n} - X_{v:n})^{-p}}$$

where k and p are chosen appropriately. It is shown that the estimator is not consistent for $k = 1$ and $p < 1$.

Basing his ideas on Grenander [5], but using more simple functionals of spacings, Venter [8] proposes two estimators of the mode θ_x which are formed on the Chernoff [2]'s idea that the mode should be the midpoint of the interval which contains the most of the observations of a unimodal continuous density function f defined on a bounded interval. Venter [8]'s estimators of $\hat{\theta}_x$ are

$$\hat{\theta}_{1x} = \frac{1}{2}(X_{K_n+r_n:n} + X_{K_n-r_n:n}) \qquad (1)$$

and

$$\hat{\theta}_{2x} = X_{K_n:n}$$

where r_n is an integer specified appropriately i.e. $r_n \approx An^v$ with A is a positive constant, $0 < v < 1$, and K_n is defined as

$$V_{K_n = \min}\{V_j : r_n + 1 \leq j \leq n - r_n\} \qquad (2)$$

where

$$V_j = X_{j+r_n:n} - X_{j-r_n:n}; \quad r_n + 1 \leq j \leq n - r_n.$$

Venter [8] shows these direct estimators of the mode are consistent in addition to their other properties.

Venter [8] as well as Grenander [5] satisfied with the asymptotic results. The exact distributional results are possible to obtain by using some results of Pyke [7] and Young [10] among others; however, the results possibly will be very complicated and restrictive. Especially the difficulties concerning the dependency structure that occurs in construction of spacings may block such results to be useful in most cases.

In this paper, Venter [8]'s mode estimation method for univariate distributions is extended to the bivariate case by using some properties of the concomitant random variables. The results of the paper are presented in the subsequent sections and an illustration is provided.

2. Estimation of the Mode

It will be convenient to recall some of the earlier results which form the basis of the result of this paper and introduce the assumptions underlying estimation process.

Probably, the most important and basic result is Theorem 1 of the Chernoff [2] which states that as $n \to \infty$ majority of observations accumulates with probability one in the center of a bounded interval that includes the true mode. So, it is reasonable to expect that the minimal spacings occur around the true mode at which majority of observaitons are most likely to be taken. Venter [8] uses this idea suggesting his consistent estimators of the mode given in (1).

There is no vital problem of extending this result to the two dimensional random vectors with unimodal bivariate distributions $F(x, y)$ having unimodal marginals.

In this setting, it is reasonable to think that the most of the bivariate observations are accumulated in a region of the support set of the bivariate distribution which contains the mode with probability one. This follows from the argument that it is possible to find two intersecting intervals one of which contains the true mode of the marginal distribution and the other one includes true mode of the conditional (or marginal) d.f. of the second r.v. conditioned on the first r.v i.e., $F_{Y/X=x}(y)$. This is accomplished by invoking the Theorem 1 of Chernoff [2] for the marginal distribution and the marginal conditional distribution. The subsequent discussion is based on this fact.

Let (X, Y) be a random vector that has absolutely continuouss d.f. $F(x, y)$. The assumptions on the probability density function f(x,y) of the random vector are as follows:

• **A1.** The marginal density function $f_X(x) > 0$ for some known constants a,b, such that $-\infty < a < x < b < \infty$ and similarly the conditional density function of the r.v Y given $X = x$ is $f_{Y/X=x}(y) > 0$ for some constants such that $-\infty < c < y < d < \infty$.

• **A2.** $f_X(x)$ and $f_{Y/X=x}(y)$ are continuous over the intervals (a,b), (c,d) respectively and they get their unique maximum values at $\theta_x \in (a,b)$ and $\theta_y \in (c,d)$

Let $(X_1,Y_1),(X_2,Y_2),(X_3,Y_3),...,(X_n,Y_n)$ be iid random vectors with the common bivariate distribution as above. If these random vectors are ordered with respect to their first component, (i.e., X), the new random vectors are represented as $(X_{1:n},Y_{[1:n]}),(X_{2:n},Y_{[2:n]}),(X_{3:n},Y_{[31:n]}),...,(X_{n:n},Y_{[n1:n]})$, where $Y_{[i:n]}$ denotes the concomitant of the i th order statistics $X_{i:n}$ which means that $Y_{[i:n]} = Y_j$ if $X_{i:n} = X_j$. Some properties of the concomitants can be found in David and Nagaraja [3] and Yang [9] An important lemma of Bhattacharya [1] that we will use in conjuction with the concomitants is the following:

Lemma. For every n and almost all $(X_1,X_2,X_3,...,X_n), Y_{[1:n]},...,Y_{[2:n]},...,Y_{[n:n]}$ are conditionally independent given $X_1, X_2, X_3,...,X_n$.

An obvious result of this Lemma is that the conditional $F(Y_{[i:n]}|x_i)$, given $X_{i:n} = x_i i = 1,2,...,n$, are iid standard uniform r.v's. They do not depend on the x_i, moreover they are also independent of $X_{i:n}$ as indicated in David and Nagaraja [3]

Let $Y_{[K_n-r_n:n]}, Y_{[K_n-r_n+1:n]}, Y_{[K_n-r_n+2:n]},...,Y_{[K_n+r_n:n]}$ be the concomitants that fall between the ordered statistics $X_{K_n-r_n:n}$ and $X_{K_n+r_n:n}$, inclusive, that constitute minimal r_n spacings in the Venter [8]'s mode estimator for the first component of the iid random vector defined as in (1) and (2). Next, define $v_n = 2r_n + 1$ and order these concomitants as $Y_{1:v_n}, Y_{2:v_n}, Y_{3:v_n},...,Y_{v_n:v_n}$. Let $\{c_n\}$ be a sequence of appropriate integers to be discussed further down the subsequent pages. Also, define the following:

$$W_i = Y_{i+c_n:v_n} - Y_{i-c_n:v_n} \; ; c_n + 1 \leq i \leq v_n - c_n \; .$$

and

$$W_{L_n} = \min\{W_i : c_n + 1 \leq i \leq v_n - c_n\} \tag{3}$$

Finally, we define two estimators for the second domponent θ_y of the vector valued mode (θ_x, θ_y) of the bivariate distribution.

$$\hat{\theta}_{1y} = \frac{1}{2}(Y_{L_n+c_n:v_n} + Y_{L_n-c_n:v_n}) \tag{4}$$

and

$$\hat{\theta}_{2y} = Y_{L_n : v_n}.$$

The first component θ_x, of the vector valued parameter is estimated using one of the estimator in (1), the second component of the parameter is estimated by using one of the estimators in (4).

The suggested estimators $\hat{\theta}_{1y}$ and $\hat{\theta}_{2y}$ are shown to be consistent in our result, which is the modified form of the Venter [8]'s Theorem 1. Before giving the statment of the result, let us give some additional definitions. In the following the notations of the Venter will be adopted. Let $\delta > 0$ be a constant.
Define,

$$\alpha(\delta) = \alpha_1(\delta)/\alpha_2(\delta)$$

where

$$\alpha_1(\delta) = \min \left\{ f(y|x) : \theta_y - \delta \le y \le \theta_y + \delta \right\}$$

and

$$\alpha_2(\delta) = \max \left\{ f(y|x) : c < y \le \theta_y - 2\delta, \ \theta_y + 2\delta \le y < d \right\}.$$

Our main result is as follows:

Theorem. If the assumptions **A1** and **A2** are satisfied and the conditions for all small enough δ

$$\alpha(\delta) > 1,$$

as $v_n \to \infty$

$$c_n / v_n \to 0$$

and for all $0 < \lambda < 1$,

$$\sum v_n \lambda^{c_n} < \infty$$

hold, then the estimators $\hat{\theta}_{1y}$ and $\hat{\theta}_{2y}$ converge to θ_y with porbability one.

Hence, the two components of the vector valued parameter (θ_x, θ_y) can be estimated by using the estimators (1) and (4). The sample size for the second component's estimation $2v_n + 1$ becomes relatively small as compared to the sample size n; however, it should be noted that the second estimator uses the information obtained in the estimation of the first component. This use of information can be thought as a compensation for the lost sample size $n - v_n$. This effect can be seen in the definiton of $\alpha(\delta)$ by writing $f(y|x)$ explicitly.

3. Proof and an Example
In this section an outline of the proof of the theorem will be given. Then, an application of the mode estimation for a bivariate distribution will be considered by using a simulated observations from a bivariate distribution.

Proof. Because of the Theorem 1 of Chernoff [2], most of the observations for the unimodal bivariate d.f.'s cumulated around a point which is common to both density functions; namely, $f_X(x)$ and $f_{Y|X=x}(y|x)$. On the other hand, it obviously is always possible to express a bivariate density function as

$$f(x,y) = f_X(x)f_{Y|X=x}(y|x).$$

A consistent mode estimator is found for the marginal density function f_X by (1). Based on this estimation, the mode estimator for the marginal conditional density function $f_{Y|X=x}(y|x)$ found by (4). This is allowed because of Lemma given above. Based on these rv's, one can construct order statistics without having difficulty by means of dependency to define c_n spacings.

The consistency part of the proof follows the same line that of the proof of Theorem 1 of Venter [8] with n replaced by v_n and r_n replaced by c_n Thus, the proof is completed.

Choice of c_n is based on the requirements $c_n/v_n \to 0$ and $\sum v_n \lambda^{c_n} < \infty$ in the theorem above; it should be $c_n \sim A v_n^v$ for a poistive constant A and $0 < v < 1$. A suitable choice is made by considering rate of convergence of the estimators to the true value of the mode. This is accomplished by Theorem 2 of Venter [9] which can be also used for the main result of this work.

For illustration, consider the set of generated pseudo observations for the random sample $(X_1, Y_1), (X_2, Y_2), (X_3, Y_3),..., (X_{300}, Y_{300})$ with a bivariate distribution having the probability density funciton given below

$$f(x,y) = \begin{cases} \dfrac{(x-y)(\frac{\pi}{2}+y)\cos(y)\sin(x-y)}{\pi^2} & , y < x < y+\pi, -\pi/2 < y < \pi/2 \\ \\ 0 & , \text{otherwise} \end{cases}$$

The random vector has the mean vector $\mu = [2.1659, 0.2976]'$ and variance covariance matrix

$$V = \begin{bmatrix} 0.7578 & 0.3789 \\ 0.3789 & 0.3788 \end{bmatrix}$$

This is a skewed bivariate distribution; so, its mean vector is different than its mode which is at the point $(2.4867, 0.4580)$. The generated data and its scatter plot presented in Table 1 and Figure 2 respectively.

Table 1: A generated data for the given bivariate (X, Y) r.v.

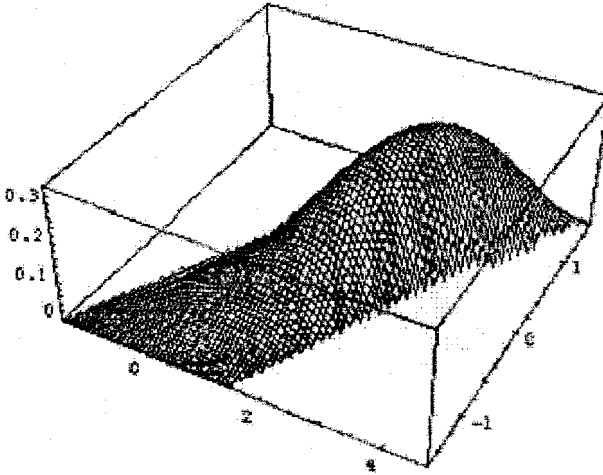|				
(4.6627,1.3750)	(1.9945,0.4245)	(2.6520,1.0151)	(3.2745,0.5033)	(2.1752,-0.4677)
(2.2617,-0.1861)	(2.7197,0.6536)	(2.0604,1.4614)	(1.8956,0.8533)	(2.5990,0.9607)
(1.6162,-0.5192)	(5.4164,-0.2591)	(2.1556,0.4545)	(3.7997,0.7737)	(3.5005,1.0681)
(2.0374,-0.5762)	(1.2333,-0.7346)	(4.1557,1.4788)	(1.6154,0.1661)	(3.1522,0.5376)
(2.2712,0.0246)	(5.0114,0.6240)	(0.9442,-0.7223)	(2.3670,0.3684)	(3.5126,0.2507)
(2.5918,0.8705)	(2.0515,0.2698)	(2.3318,0.1441)	(3.4525,0.4025)	(1.8245,-0.3301)
(2.0176,0.2601)	(2.5955,0.7228)	(1.6609,0.6606)	(2.5647,1.4760)	(3.6684,0.3948)
(2.9003,0.3572)	(1.7400,-0.5002)	(0.9751,-0.1519)	(1.2614,0.5862)	(0.9351,0.0745)
(0.4585,-0.7291)	(1.2263,-0.1412)	(2.5135,-0.1692)	(3.1196,0.9874)	(2.5671,0.7349)
(2.5027,0.4814)	(1.5046,-0.3162)	(.9381,0.0345)	(2.1863,1.2584)	(2.0423,0.6437)
(1.3096,0.7114)	(2.9604,1.3157)	(2.0792,0.1747)	(2.4599,0.5291)	(1.9535,-0.1513)
(2.2557,-0.2109)	(2.4944,0.7345)	(2.2406,0.5362)	(3.3570,0.7055)	(1.6365,0.6376)
(2.8314,0.5201)	(1.5570,0.0666)	(1.4410,-0.6497)	(2.188,0.0650)	(2.4542,0.4184)
(0.1771,-0.4946)	(1.6726,0.9301)	(2.0408,0.5288)	(1.9939,0.9121)	(2.4587,0.3934)
(2.0653,0.6955)	(1.2099,-0.5628)	(1.7377,1.0155)	(1.8573,0.7737)	(3.5242,1.1912)
(2.4762,-0.2161)	(2.2365,0.2664)	(2.2479,-0.4345)	(3.4744,0.6624)	(1.4502,1.5039)
(1.8188,0.7626)	(2.3078,0.3701)	(2.1760,1.5459)	(2.6117,-0.0677)	(1.2431,-0.4514)
(1.0549,-0.5925)	(2.4162,1.1181)	(3.6292,1.2035)	(2.9314,0.4743)	(0.5504,-0.8884)
(0.6295,-0.6314)	(2.1245,0.2525)	(5.0269,0.7996)	(3.5053,0.3514)	(2.2513,0.3115)
(1.1626,-0.9760)	(3.2594,1.0760)	(2.1462,0.7746)	(3.4185,0.3272)	(0.7446,-0.7100)
(1.2706,0.2157)	(2.1203,0.5308)	(1.6676,0.6656)	(0.7936,-0.6995)	(3.7477,0.6723)
(1.8736,-0.5817)	(2.9407,0.9363)	(2.5108,0.6279)	(2.5485,0.4436)	(2.2006,1.2307)
(3.5679,1.1556)	(2.8441,0.9473)	(3.8269,1.3652)	(2.6607,1.1214)	(3.1653,0.5591)
(0.4431,0.1465)	(0.8775,-0.2195)	(2.2651,0.1780)	(1.0699,0.1776)	(1.9961,0.5067)
(1.6745,0.0151)	(2.3676,0.6163)	(2.0659,1.1377)	(1.2509,0.7729)	(2.7356,1.0779)
(0.2214,-0.7815)	(0.1355,-0.5346)	(2.9056,1.5973)	(1.4001,-0.3677)	(2.2703,1.2629)
(0.8321,-0.0841)	(2.1633,0.2362)	(2.9178,0.8720)	(1.2696,1.0607)	(1.1451,0.0734)
(1.8692,0.1077)	(3.3857,0.6069)	(1.2376,-0.5426)	(2.4401,0.5571)	(3.1099,1.3621)
(2.7033,0.2284)	(3.0845,0.4688)	(1.4033,-0.5518)	(3.0680,1.0135)	(1.1121,0.8058)
(0.7184,-0.9079)	(1.8235,0.2991)	(1.4001,0.9031)	(2.7188,-0.3090)	(2.7960,1.3690)
(1.1235,-0.6493)	(3.2338,0.9551)	(1.1275,-0.9759)	(1.6096,-0.5459)	(2.0745,0.8596)
(5.0172,0.5417)	(2.7083,0.1265)	(0.3606,-1.1001)	(1.6002,0.2362)	(1.8778,-0.1536)
(3.4636,0.7346)	(3.5401,1.1136)	(3.1624,1.1618)	(3.3947,-0.1081)	(0.9609,-0.0610)
(0.3884,-0.2231)	(1.4586,0.9340)	(1.9028,0.1867)	(2.0460,-0.5288)	(1.9125,0.2966)
(2.5699,0.8886)	(2.2046,0.4818)	(5.5674,1.2489)	(1.5242,-0.2477)	(1.8923,0.4548)
(2.9901,0.7848)	(2.4235,1.1764)	(1.8290,-0.3097)	(3.1876,0.6137)	(2.1241,-0.4974)
(1.9709,0.4138)	(1.0683,-0.3385)	(2.1000,0.0158)	(2.8943,1.0695)	(2.6122,0.5071)
(1.7003,-0.1156)	(0.9952,0.1469)	(2.3698,0.5493)	(1.6787,0.3853)	(0.9605,-0.7314)
(1.2127,0.4468)	(0.6046,-0.2041)	(2.0053,0.0725)	(3.3545,1.0017)	(3.8223,1.1463)
(2.7496,0.8765)	(2.4644,0.1622)	(2.1976,0.4064)	(1.6676,0.0632)	(1.2584,0.1923)
(3.8614,0.5262)	(3.8661,1.3986)	(1.7267,0.6687)	(1.8993,0.1594)	(3.5855,0.7867)
(2.1766,0.6032)	(2.8884,1.0056)	(1.1885,-0.5111)	(0.8826,-0.4079)	(1.9466,0.0600)
(1.3633,0.7280)	(2.2519,0.5469)	(3.1895,-0.3692)	(2.4379,0.5210)	(1.9286,0.9667)
(2.6027,0.7117)	(0.5209,0.1103)	(2.0920,1.1265)	(2.1898,0.5888)	(0.5748,-1.1074)
(1.2224,-0.5743)	(1.4635,0.9089)	(1.2568,-0.0954)	(1.7988,0.8165)	(2.7527,0.7720)
(1.7114,0.1690)	(3.0456,1.4985)	(1.8065,-0.5018)	(1.4456,0.2150)	(1.0456,-0.0068)
(2.1668,0.5114)	(0.0706,-0.6865)	(3.4882,1.0685)	(3.6439,1.0264)	(2.8841,0.7809)
(0.3922,-0.6290)	(2.5916,0.1665)	(2.3156,0.0611)	(2.2735,-0.3478)	(0.7362,-0.4486)
(3.9879,1.5953)	(2.1062,0.0872)	(1.5963,-0.1032)	(1.5076,-0.4653)	(2.2329,0.7104)
(3.7143,1.0774)	(2.4154,0.8452)	(1.5421,0.8555)	(1.1708,-0.5446)	(0.6585,0.3609)
(1.7908,0.4102)	(0.4593,-0.0203)	(2.4668,0.5639)	(1.8651,-0.0054)	(2.1236,-0.3626)
(5.5980,1.2656)	(2.2196,0.5611)	(1.7225,0.7525)	(2.3501,-0.1735)	(1.7041,0.4910)
(1.6295,0.3146)	(2.8539,0.2643)	(5.6212,0.6376)	(2.6098,0.1055)	(2.5643,0.7135)
(2.6586,0.6565)	(1.0017,-0.0210)	(2.3080,1.1287)	(3.4046,0.4942)	(1.2496,0.1506)
(3.7760,1.1215)	(0.6457,0.1965)	(1.6627,0.5646)	(0.3229,-0.5947)	(0.8545,0.0355)
(3.8039,1.1060)	(1.4505,-0.1035)	(3.0513,0.8465)	(2.1781,0.1868)	(2.3059,-0.6665)
(2.5456,0.5508)	(5.0421,0.7179)	(3.4287,1.0473)	(1.9165,-0.6677)	(1.8770,0.3590)
(1.8589,0.6683)	(0.8003,-0.2160)	(2.2658,1.0946)	(1.8644,0.5494)	(2.0308,-0.4273)
(2.1402,0.2170)	(1.8068,-0.2876)	(1.2394,-0.5911)	(2.0904,-0.1826)	(2.1594,0.0327)
(0.6484,-0.5910)	(1.7734,0.8920)	(1.1111,-0.0646)	(2.6093,0.4094)	(1.8971,-0.2391)

Figure 1: The probability density function of the bivariate random vector used in the example.

First, fix $r_n = n^{1/2} \approx 17$ for the given data. Of course better choices for r_n are possible to increase the rate of convergence of the estimator to the true mode as

much as the conditions in Theorem 1 of Venter [9] is satisfied namely for a positive constant A and $0 < v < 1$, $r_n \approx An^v$. We find that

$$V_{K_n} = \min\{V_j : r_n + 1 \le j \le n - r_n\}$$

$$= x_{172:300} - x_{138:300} = 0.1943$$

for $K_n = 155$. If the first or the second estimator in (1) is used we find that

$$\hat{\theta}_{1x} = \frac{1}{2}(x_{172:300} + x_{138:300}) = (2.22735 + 2.0792)/2 = 2.1764 \qquad (5)$$

or

$$\hat{\theta}_{2x} = x_{155:300} = 2.1751.$$

Now, after reordering the observed concomitants

$$y_{[138:300]}, y_{[139:300]}, y_{[140:300]}, \cdots, y_{[172:300]}$$

Between $x_{138:300}$ and $x_{172:300}$ and taking $c_n = v_n^{1/2} \approx 6$, with $L_n = 21$

$$W_{L_n} = \min\{W_i : c_n + 1 \le i \le v_n - c_n\}$$

$$= y_{27:35} - y_{15:35} = 0.3374$$

is found. By using the estimators defined in (4) the second component of the estimation is obtained as

$$\hat{\theta}_{1y} = \frac{1}{2}(y_{27:35} + y_{15:35}) = 0.3802 \tag{6}$$
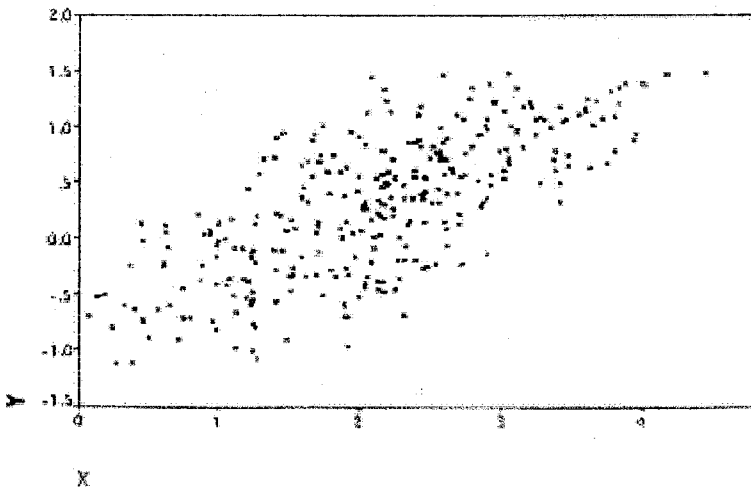
or

$$\hat{\theta}_{2y} = y_{21:35} = 0.3611.$$

Figure 2: The scatter plot for the generated bivariate data.

As a result, (2.1764, 0.3802) or (2.1751, 0.3611) can be considered as an estimation of the mode which was actually (2.4867, 0.4580).

If we had chosen the r.v. Y to order the random vector (X,Y) with respect to, we would find the estimate little different by using the estimators $(\hat{\theta}_{1x}, \hat{\theta}_{1y})$ and $(\hat{\theta}_{2x}, \hat{\theta}_{2y})$ respectively. It may be helpful that to have some information about the underlying distribution to decide which component to choose first for ordering the random vector in order to increase the performance of the estimator. The other possible way is to take their componentwise artihmetic mean.

The result of this paper can be extended for the higher dimensional random vectors in a nested way. However, as dimension increases we will need larger sample sizes.

REFERENCES
[1]   Bhattacharýa P.K. (1974). Convergence of sample paths of normalized sums of induced order statistics. Ann. Statist., 2, 1034-1039.
[2]   Chernoff, H. (1964). Estimation of the mode. Ann. Inst. Statist. Math., 16,36-41.
[3]   David, H.A. and Nagaraja, H.N. (1998). Concomitants of order statistics. Handbook of Statistics, Vol. 16, eds. N. Balakrishnan, C.R. Rao. 487-513.
[4]   Feller, W. (1971). An Introduction to Probility Theory and its Applications. Volume II, 2nd ed. Wiley, New York.
[5]   Grenander, U. (1965). Some direct estimates of the mode. Ann. Math. Statist., 36,131-138.
[6]   Parzen, E. (1962). On estimation of a probability density function and mode.Ann. Math. Statist., 33, 1065-1076.
[7]   Pyke, R. (1965). Spacings. J. Roy. Statist. Soc., Ser. B, 27, 395-496.
[8]   Venter, J.H. (1967). On estimation of the mode. Ann. Math. Statist., 38, 1447-1455.
[9]   Yang, S.S. (1977). General distribution theory of the concomitants of the order statistics. Ann. Statist., 5,996-1002.
[10] Young, D.H. (1967). Recurrence relations between the p.d.f.'s of order statistics of dependent variables, and some applications. Biometrika, 54, 283-292.