Inference in a Curved Binomial Distribution

Sanchayita Sadhu[*], Babulal Seal[†]

[*]Department of Statistics, Burdwan University, West Bengal, INDIA.

[†]Associate Professor, Department of Statistics, Burdwan University, West Bengal, INDIA.

Corresponding Addresses:

[*]sanchayita.sadhu@gmail.com, [†]babulal_seal@yahoo.com

Research Article

Introduction

Curved exponential families of distributions play an important role in theory of statistics. In case of inference, for curved exponential family we can not find test statistic as good, powerful and efficient as the exponential family. But there is a ray of hope due to Bradly Efron. He has shown in his paper, "Defining the Curvature of a Statistical Problem (with Application to Second Order Efficiency)" [1] that families with small curvature enjoy the good statistical properties of exponential families.

In curved exponential family, variance of maximum likelihood estimator (MLE) exceeds the Cramer-Rao lower bound. For curved exponential family MLE is not the sufficient statistic in general. Nonetheless, with small curvature, curved exponential family enjoys good statistical properties. Firstly it is to be found that the values of the involved parameters of the distributions, for which, curvature have small values. Hence, after finding such values, it can be suggested that for the values of the parameter, the test statistic of certain hypothesis is equivalent to that of the corresponding exponential family. In this case it is found that some inference procedures are available fragmentally for continuous distributions. But for discrete cases we don't find any procedure. Motivated by my (Sanchayita Sadhu) project work (submitted during the period of my post graduate study), this paper wants to find some techniques of inference in binomial distribution – a discrete distribution.

Curved Exponential Family:-

Let $X = (X_1, X_2,\ldots, X_d)$ have a distribution P_θ, $\theta \in \Theta \subseteq R^q$. Suppose P_θ has a density (pmf) of the form

$$f\left(x \mid \theta\right) = \exp(\sum_{i=1}^{k} \eta_i(\theta) T_i(x) - \Psi(\theta)) h(x),$$

where k > q. Then the family $\{P_\theta, \theta \in \Theta\}$ is called curved exponential family.

Some Examples of Curved Exponential Family:

1. A set of independently and identically distributed random variables which follow $N(\theta, \theta^2)$, θ is the unknown parameter involved in the distribution [4].

2. A random variable, say, X that follows gamma $(\alpha, 1/\alpha)$, α is the unknown parameter involved in the distribution.

3. X and Y, say, are two random variables such that
$$\begin{matrix} X \sim B(n, p) \\ Y \sim B(m, p^2) \end{matrix} \succ independently$$
where p is the unknown parameter involved in both the distributions [4].

4. Let a random variable, say, $X = (X_1, X_2)$ has a Bivariate normal distribution with zero mean, standard deviation equal to 1 and correlation parameter ρ, ρ is the involved unknown parameter [4].

5. Suppose $Z_i = z_i$, i = 1,2,…,n. X_i are independent Poi(λz_i) variables and $Z_1, Z_2,\ldots, Z_n$ have some joint p.m.f $p(z_1, z_2,\ldots, z_n)$. It is implicitly assumed that each $Z_i > 0$ with probability 1. Then the joint p.m.f of $(X_1, X_2,\ldots, X_n, Z_1, Z_2,\ldots, Z_n)$ is

$$f(x_1, x_2,\ldots, x_n, z_1, z_2,\ldots, z_n \mid \lambda) = e^{-\lambda\sum_{i=1}^{n} z_i + (\sum_{i=1}^{n} x_i)\log\lambda} \prod_{i=1}^{n} \frac{z_i^{x_i}}{x_i!} p(z_1, z_2,\ldots, z_n)$$
$$I x_1,\ldots, x_n \in N_0,$$
$$I z_1,\ldots, z_n \in N_1$$

N_0 = set of non-negative integer, N_1 = set of positive integers [4].

6. Equicorrelation Multivariate Normal distribution: suppose $(X_1, X_2,\ldots, X_n)$ are jointly multivariate normal with general mean μ_i, variances all 1 and a common pair wise correlation ρ. This is an example of curved exponential family [4].

Applications of curved exponential family:-

First of all, some real life data, in which the random variables belong to curved exponential family, is

discussed. The analysis, done in this paper, hope to be applied to the data to draw certain inference.

Some practical examples of curved exponential family are discussed here.

➢Mixed Ancestral Graph (MAG):-

The family of distributions represented by a linear MAG M over a set of k variables is a locally parameterized curved exponential family of dimension equal to k(k+1)/2 minus the number of pairs of variables in M that are not adjacent to each other. (Theorem 4; Parameterizing and Scoring Mixed Ancestral Graphs by Thomas Richdson & Peter Spirtes [2]).

➢Lazega's Lawyer dataset:-

Lazega's Lawyer dataset [5] is another example where the random variables of the dataset are from curved family.

➢Use in Social Networking:-

David. R. Hunter, in his paper, "Curved Exponential Family Models for Social Networks", [3] has stated on the usefulness of curved exponential family models on generalization of exponential random graph model (ERGMs).

Concept of Statistical Curvature:-

The concept of mathematical curvature extends to the curved lines in Euclidian k-space, E^k, say, £= {η_θ, $\theta \in \Theta$}, where Θ is the interval of real line. For each θ, η_θ is a vector in E^k whose component wise derivatives with respect to θ is denoted by

$$\dot{\eta}_\theta \equiv \left(\frac{\partial}{\partial\theta}\right)\eta_\theta \quad \& \quad \ddot{\eta}_\theta \equiv \left(\frac{\partial^2}{\partial\theta^2}\right)\eta_\theta$$

These derivatives are assumed to exist continuously in neighborhood of a value of θ where it is wished to define the curvature. Let us also suppose that a k×k symmetric non-negative definite matrix Σ_θ is defined continuously in θ.

Let M_θ be a 2×2 matrix, with entries denoted

$$\nu_{20}(\theta), \nu_{11}(\theta), \nu_{02}(\theta),$$

defined by

$$M_\theta \equiv \begin{pmatrix} \nu_{20}(\theta) & \nu_{11}(\theta) \\ \nu_{11}(\theta) & \nu_{02}(\theta) \end{pmatrix} \equiv \begin{pmatrix} \dot{\eta}_\theta'\Sigma_\theta\dot{\eta}_\theta & \dot{\eta}_\theta'\Sigma_\theta\ddot{\eta}_\theta \\ \ddot{\eta}_\theta'\Sigma_\theta\dot{\eta}_\theta & \ddot{\eta}_\theta'\Sigma_\theta\ddot{\eta}_\theta \end{pmatrix} \quad (1)$$

and let

$$\gamma_\theta \equiv \left(|M_\theta|/\nu_{20}^3(\theta)\right)^{1/2} \quad (2)$$

Then γ_θ is "the curvature of £ at θ with respect to inner product Σ_θ".

Statistical Curvature:-

γ_θ (given in (2)), the statistical curvature of F at θ, is the geometric curvature of £={$\eta_\theta : \theta \in \Theta$} at θ with respect

to the covariance inner product Σ_θ as defined in (1) and (2).

Here F will stand for the family of densities {$f_\theta(x): \theta \in \Theta$}, our curved exponential family.

Procedure for Finding Curvature of Various Distributions from Curved Exponential Family:-

First the pmf/pdf of the random variables is arranged as the known form of curved exponential family. i.e. of the form

$$f(x|\theta) = \exp\left(\sum_{i=1}^{k}\eta_i(\theta)T_i(x) - \Psi(\theta)\right)h(x)\cdot$$

Then finding the values of $\eta_i(\theta)$'s and $T_i(x)$'s the matrix

$$D\begin{pmatrix} T_1 \\ : \\ T_i \end{pmatrix}$$

is constructed, $\dot{\eta}_\theta'\Sigma_\theta\dot{\eta}_\theta$, $\dot{\eta}_\theta'\Sigma_\theta\ddot{\eta}_\theta$, $\ddot{\eta}_\theta'\Sigma_\theta\dot{\eta}_\theta$ and $\ddot{\eta}_\theta'\Sigma_\theta\ddot{\eta}_\theta$ are computed and the matrix $M_\theta = \begin{pmatrix} \dot{\eta}_\theta'\Sigma_\theta\dot{\eta}_\theta & \dot{\eta}_\theta'\Sigma_\theta\ddot{\eta}_\theta \\ \ddot{\eta}_\theta'\Sigma_\theta\dot{\eta}_\theta & \ddot{\eta}_\theta'\Sigma_\theta\ddot{\eta}_\theta \end{pmatrix}$ is

constructed. $|M_\theta|/\nu_{20}^3(\theta) \equiv |M_\theta|/(\dot{\eta}_\theta'\Sigma_\theta\dot{\eta}_\theta)^3$ is then computed. If there are some negative values in the above result, then the absolute values are taken. Finally $\left(|M_\theta|/\nu_{20}^3(\theta)\right)^{1/2}$ i.e. the statistical curvature for various values of θ is calculated.

Using this method the values of the statistical curvature corresponding to the possible values of the involved parameters are found. From these values it can be easily concluded that for which values of the involved parameters the curvature has small values and it can be said that for these values of the parameters, the given curved exponential family enjoys the good statistical properties of exponential families.

The Area of Work:-

The details work of analysis to find these values of involved parameters for which the curvature has small values is done on the following distributions which are some particular examples of curved exponential family. The examples are given below:

1. A set of independently and identically distributed random variables which follow N (θ, θ^2), θ is the unknown parameter involved in the distribution.
2. A random variable, say, X that follows gamma($\alpha, 1/\alpha$), α is the unknown parameter involved in the distribution.
3. X and Y, say, are two random variables such that
$$\left.\begin{array}{l} X \sim B(n, p) \\ Y \sim B(m, p^2) \end{array}\right\} \text{independently}$$
where p is the unknown parameter involved in both the distributions.

4. Let a random variable, say, $X = (X_1,X_2)$ has a Bivariate normal distribution with zero mean, standard deviation equal to 1 and correlation parameter ρ, ρ is the involved unknown parameter.

Curvature of some curved distributions:-

1. *On Normal Distribution:-*

Let $X_1,X_2,…,Xn$ are iid $N(\theta,\theta^2)$.

Here

$$f(x|\theta) = \left(\frac{1}{\theta\sqrt{2\pi}}\right)^n \exp\left(-\frac{1}{2\theta^2}\sum_i (x_i - \theta)^2\right)$$

$$= \left(\frac{1}{\theta\sqrt{2\pi}}\right)^n \exp\left(-\frac{1}{2}\sum_i x_i^2\Big/\theta^2 + \frac{1}{2}\sum_i x_i\Big/\theta - n\Big/2\right)$$

$$= k(\theta)\exp(\eta_1(\theta) T_1 + \eta_2(\theta) T_2) h(x)$$

Here,

$\eta_1(\theta) = 1/\theta^2$, $\eta_2 = 1/\theta$.

$T_1 = \sum_{i=1}^{n} x_i^2$, $T_2 = \sum_{i=1}^{n} x_i$.

Now, the work is to find $V_{(\sum_i x_i^2)}$, $V\left(\sum_i x_i\right)$ and $\text{cov}\left(\sum_i x_i^2, \sum_i x_i\right)$.

It is known that the moment generating function of X will be

$$M_X(t) = \exp\left(\theta t + \frac{1}{2}\theta^2 t^2\right) \quad (1.0)$$

Now,

$$V\left(\sum_i x_i^2\right) = \sum_i V\left(x_i^2\right) = E\left(x_i^4\right) - \left[E\left(x_i^2\right)\right]^2$$

Therefore from (1.0) it can be computed that

$E(X^4) = 10\theta^4$

$\therefore V(X_i^2) = E(X_i^4) - [E(X_i^2)]^2 = 6\theta^4$

$$\therefore V\left(\sum_i X_i^2\right) = \sum_i V\left(X_i^2\right) = 6n\theta^4$$

$$V\left(\sum_i X_i\right) = \sum_{i=1}^{n} V(X_i) = n\theta^2.$$

Now, $\text{cov}\left(\sum_i X_i^2, \sum_i X_i\right) = \sum_{i=1}^{n}\text{cov}\left(X_i^2, X_i\right)$

$\text{Cov}(X_i^2, X_i) = E(X_i^3) - E(X_i^2)E(X_i) = 2\theta^3$

$\therefore \text{cov}\left(\sum_i X_i^2, \sum_i X_i\right) = 2n\theta^3.$

Hence the dispersion matrix (variance-covariance matrix) is given by

$$\Sigma_\theta = \begin{pmatrix} 6n\theta^4 & 2n\theta^3 \\ 2n\theta^3 & n\theta^2 \end{pmatrix}.$$

Also,

$$\eta_\theta = \begin{pmatrix} \dfrac{1}{\theta^2} \\ \dfrac{1}{\theta} \end{pmatrix}$$

$$\therefore \dot{\eta}_\theta = \begin{pmatrix} \dfrac{-1}{\theta^3} \\ \dfrac{-1}{\theta^2} \end{pmatrix} \quad \& \quad \ddot{\eta}_\theta = \begin{pmatrix} \dfrac{3}{\theta^4} \\ \dfrac{2}{\theta^3} \end{pmatrix}$$

$$\therefore \dot{\eta}_\theta{}'\Sigma_\theta \dot{\eta}_\theta = \begin{pmatrix} \dfrac{-1}{\theta^3} & \dfrac{-1}{\theta^2} \end{pmatrix} \begin{pmatrix} 6n\theta^4 & 2n\theta^3 \\ 2n\theta^3 & n\theta^2 \end{pmatrix} \begin{pmatrix} \dfrac{1}{\theta^2} \\ \dfrac{1}{\theta} \end{pmatrix}$$

$$= \frac{11n}{\theta^2} \quad (1.1)$$

Similarly,

$$\dot{\eta}_\theta{}'\Sigma_\theta \ddot{\eta}_\theta = -\frac{30n}{\theta^3}. \quad (1.2)$$

$$\ddot{\eta}_\theta{}'\Sigma_\theta \dot{\eta}_\theta = -\frac{30n}{\theta^3} \quad (1.3)$$

$$\ddot{\eta}_\theta{}'\Sigma_\theta \ddot{\eta}_\theta = \frac{82n}{\theta^4} \quad (1.4)$$

So, M_θ is given by

$$M_\theta = \begin{pmatrix} \dfrac{11n}{\theta^2} & \dfrac{-30n}{\theta^3} \\ \dfrac{-30n}{\theta^3} & \dfrac{82n}{\theta^4} \end{pmatrix} \quad (1.5)$$

$$\therefore |M_\theta| = \begin{vmatrix} \dfrac{11n}{\theta^2} & -\dfrac{30n}{\theta^3} \\ -\dfrac{30n}{\theta^3} & \dfrac{82n}{\theta^4} \end{vmatrix} = \frac{2n^2}{\theta^6} \quad (1.6)$$

and

$$\upsilon_{20}{}^3(\theta) = \left(\frac{11n}{\theta^2}\right)^3 = \frac{1331n^3}{\theta^6} \quad (1.7)$$

$$\therefore \text{Curvature} = \gamma_\theta \equiv \left(|M_\theta|\Big/\upsilon_{20}{}^3(\theta)\right)^{1/2} = \left(\frac{2}{1331n}\right)^{1/2} \quad (1.8)$$

So, for all possible values of θ, the value of the curvature is very small.

2. *On Gamma Distribution:-*

Let X be a random variable which follows Gamma$(\alpha,1/\alpha)$.

i.e $X \sim G(\alpha,1/\alpha)$.

Then the pdf of X is given by

$$f(x|\theta) = \frac{\left(\dfrac{1}{\alpha}\right)^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\frac{x}{\alpha}} \quad (2.1)$$

$$\therefore f(x|\theta) = \frac{e^{-\frac{x}{\alpha}}}{(\Gamma(\alpha))}x^{\alpha-1}\frac{1}{\alpha}$$

$$= k(\theta)\, e^{(\alpha-1)\log x - \frac{1}{\alpha}x}\, h(x)$$

$= k(\theta)\ e^{\eta_1(\theta)T_1 + \eta_2(\theta)T_2}\, h(x)$ (2.2)

Here $\eta_1(\theta) = \alpha-1$ and $\eta_2(\theta) = 1/\alpha$

$T_1 = \log x$ and $T_2 = x$

Now to find the required variances and covariance a simulation procedure is to be followed. After computing the required expressions, using an R program the values of curvature are found.

The curvature of the mentioned Gamma distribution is given in the following figure:

Figure1

Plot of the curvature of curved Gamma distribution for various values of alpha
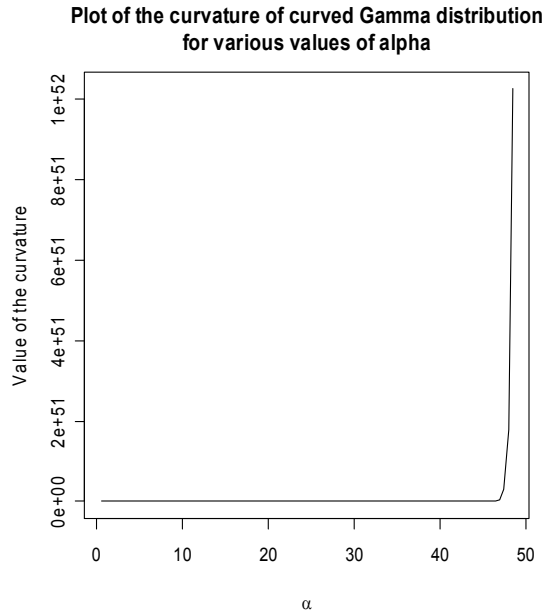
So, from Figure1, it may be concluded that for $\alpha = 0$ to 46 the value of the curvature is small.

3. *On Binomial Distribution:-*
Let X follows Binomial(n,p) and Y follows Binomial(m,p^2)
i.e

$$X \sim B(n,p)$$
$$Y \sim B(m,p^2)\ \succ independently$$

Then the joint p.m.f of (X, Y) is

$$p(x,y) = \binom{n}{x}\binom{m}{y} p^x\, p^{2y}\, (1-p)^{n-x}\, (1-p^2)^{m-y}$$

$$= \binom{n}{x}\binom{m}{y} p^{x+2y}\, (1-p)^{n-x+m-y}\, (1+p)^{m-y}$$

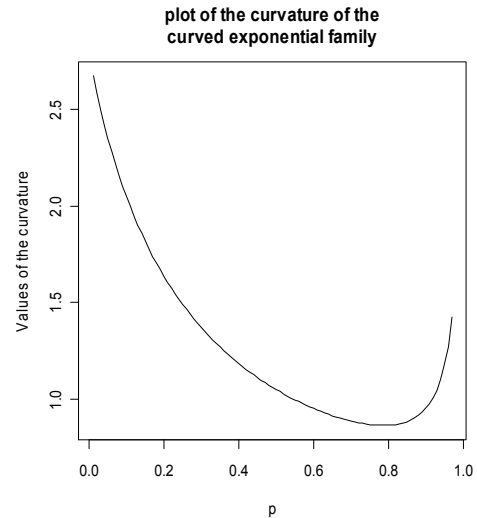$$= \exp\left[x\log\frac{p}{1-p} + y\log\frac{p^2}{1-p^2} + \{(n+m)\log(1-p) + m\log(1+p)\}\right] h(x)$$

$\therefore\ \eta_1(p) = \log\dfrac{p}{1-p}$ and $\eta_2(p) = \log\dfrac{p^2}{1-p^2}$

$T_1(x,y) = x$, $T_2(x,y) = y$.

With this information an R program is used to find the values of the curvature.

The curvature of the mentioned distribution is given in the following figure:

Figure 2

plot of the curvature of the curved exponential family

It is known that there is test statistic for single binomial distribution. Here is the joint distribution of two independent random variables which follow Bin(n,p) and Bin(m,p^2) respectively. So, for the values of p for which the value of the curvature is small, any test statistic which is valid in corresponding exponential family can not be referred.

Inference of the Above Binomial Distributions:
➤Test Procedure:
To overcome this problem this paper would like to find out a test procedure to draw an inference about the null hypothesis. Here is the likelihood ratio test.

From Figure 2 it is seen that for p=0.6 to 0.85 the value of the curvature is small.

This paper wants to test

$H_0 : p = 0.8$ vs. $H_1 : p \neq 0.8$

To find LRT firstly the maximum likelihood estimator of p is found.

$$\frac{dp(x,y)}{dp} = \binom{n}{x}\binom{m}{y} p^{x+2y-1}(1-p)^{n-x+m-y-1}(1+p)^{m-y-1}[x+2y+px-pn-p^2n-2mp^2]$$

$$\therefore \frac{d^2 p(x,y)}{dp^2} = (x-n-2pn-4pm)\binom{n}{x}\binom{m}{y}p^{x+2y-1}(1-p)^{n-x+m-y-1}(1+p)^{m-y-1} + (x+2y+px-pn-$$

$$p^2 n - 2mp^2)\binom{n}{x}\binom{m}{y}p^{x+2y-2}(1-p)^{n-x+m-y-2}(1+p)^{m-y-2}[(1-p)(1+p)(x+2y-1)-$$

$$(n-x+m-y-1)p(1+p)+(m-y-1)p(1-p)]$$

Solving $\frac{dp(x,y)}{dp}=0$, taking n=5, m=3; [Using R],

$\hat{p}_{max} = 0.9090926$ (3.0)

Therefore, for LRT

$$\lambda = \frac{\underset{H_0}{Sup}\, p(x,y)}{\underset{H}{Sup}\, p(x,y)} = \frac{\binom{n}{x}\binom{m}{y}(0.8)^{x+2y}(1-0.8)^{n-x+m-y}(1+0.8)^{m-y}}{\binom{n}{x}\binom{m}{y}(0.9090926)^{x+2y}(1-0.9090926)^{n-x+m-y}(1+0.9090926)^{m-y}}$$

Using R the values of λ will be found.

Let us consider the following table, from which a decision can be made:

Table 3

Ob.No.	(x,y)	λ	Sort(λ) $\equiv$ q	Corresponding (x,y)	Probability sum $\equiv$ s
1	(0,0)	460.0262373	0.2450758	(5,3)	0.08589935
2	(0,1)	171.7394474	0.6127021	(4,3)	0.10737418
3	(0,2)	64.1146861	0.6564672	(5,2)	0.15569256
4	(0,3)	23.9356364	1.5317867	(3,3)	0.16106127
5	(1,0)	184.0067301	1.6412016	(4,2)	0.17314087
6	(1,1)	68.6943735	1.7584320	(5,1)	0.20031996
7	(1,2)	25.6453497	3.8295451	(2,3)	0.20166214
8	(1,3)	9.5740587	4.1030881	(3,2)	0.20468204
9	(2,0)	73.6011862	4.3961701	(4,1)	0.21147681
10	(2,1)	27.4771872	4.7101868	(5,0)	0.22676505
11	(2,2)	10.2579300	9.5740587	(1,3)	0.22710059
12	(2,3)	3.8295451	10.2579300	(2,2)	0.22785556
13	(3,0)	29.4398721	10.9906500	(3,1)	0.22955426
14	(3,1)	10.9906500	11.7757079	(4,0)	0.23337632
15	(3,2)	4.1030881	23.9356364	(0,3)	0.23346020
16	(3,3)	1.5317867	25.6453497	(1,2)	0.23364895
17	(4,0)	11.7757079	27.4771872	(2,1)	0.23407362
18	(4,1)	4.3961701	29.4398721	(3,0)	0.23502914
19	(4,2)	1.6412016	64.1146861	(0,2)	0.23507632
20	(4,3)	0.6127021	68.6943735	(1,1)	0.23518249
21	(5,0)	4.7101868	73.6011862	(2,0)	0.23542137
22	(5,1)	1.7584320	171.7394474	(0,1)	0.23544791
23	(5,2)	0.6564672	184.0067301	(1,0)	0.23550763
24	(5,3)	0.2450758	460.0262373	(0,0)	0.23552256

With the help of the Table 3 the decision reached is as follows:

Table 4

If α	Reject H_0 if (x,y) $\equiv$
< 0.08	(5,3)
= 0.107	(4,3) & (5,3)
=0.15	(5,2), (4,3) & (5,3)
= 0.16	(3,3), (5,2), (4,3) & (5,3)
= 0.17	(4,2), (3,3), (5,2), (4,3) & (5,3)
= 0.200	(5,1), (4,2), (3,3), (5,2), (4,3) & (5,3)

= 0.201	(2,3), (5,1), (4,2), (3,3), (5,2), (4,3) & (5,3)
= 0.204	(3,2), (2,3), (5,1), (4,2), (3,3), (5,2), (4,3) & (5,3)
= 0.21	(4,1), (3,2), (2,3), (5,1), (4,2), (3,3), (5,2), (4,3) & (5,3)
= 0.226	(5,0), (4,1), (3,2), (2,3), (5,1), (4,2), (3,3), (5,2), (4,3) & (5,3)
= 0.2271	(1,3), (5,0), (4,1), (3,2), (2,3), (5,1), (4,2), (3,3), (5,2), (4,3) & (5,3)
= 0.2278	(2,2), (1,3), (5,0), (4,1), (3,2), (2,3), (5,1), (4,2), (3,3), (5,2), (4,3) & (5,3)
= 0.229	(3,1), (2,2), (1,3), (5,0), (4,1), (3,2), (2,3), (5,1), (4,2), (3,3), (5,2), (4,3) & (5,3)
= 0.2333	(4,0), (3,1), (2,2), (1,3), (5,0), (4,1), (3,2), (2,3), (5,1), (4,2), (3,3), (5,2), (4,3) & (5,3)
= 0.2334	(0,3), (4,0), (3,1), (2,2), (1,3), (5,0), (4,1), (3,2), (2,3), (5,1), (4,2), (3,3), (5,2), (4,3) & (5,3)
= 0.2336	(1,2), (0,3), (4,0), (3,1), (2,2), (1,3), (5,0), (4,1), (3,2), (2,3), (5,1), (4,2), (3,3), (5,2), (4,3) & (5,3)
= 0.234	(2,1), (1,2), (0,3), (4,0), (3,1), (2,2), (1,3), (5,0), (4,1), (3,2), (2,3), (5,1), (4,2), (3,3), (5,2), (4,3) & (5,3)
= 0.23502	(3,0), (2,1), (1,2), (0,3), (4,0), (3,1), (2,2), (1,3), (5,0), (4,1), (3,2), (2,3), (5,1), (4,2), (3,3), (5,2), (4,3) & (5,3)
= 0.23507	(0,2), (3,0), (2,1), (1,2), (0,3), (4,0), (3,1), (2,2), (1,3), (5,0), (4,1), (3,2), (2,3), (5,1), (4,2), (3,3), (5,2), (4,3) & (5,3)
= 0.2351	(1,1), (0,2), (3,0), (2,1), (1,2), (0,3), (4,0), (3,1), (2,2), (1,3), (5,0), (4,1), (3,2), (2,3), (5,1), (4,2), (3,3), (5,2), (4,3) & (5,3)
= 0.23542	(2,0), (1,1), (0,2), (3,0), (2,1), (1,2), (0,3), (4,0), (3,1), (2,2), (1,3), (5,0), (4,1), (3,2), (2,3), (5,1), (4,2), (3,3), (5,2), (4,3) & (5,3)
= 0.23544	(0,1), (2,0), (1,1), (0,2), (3,0), (2,1), (1,2), (0,3), (4,0), (3,1), (2,2), (1,3), (5,0), (4,1), (3,2), (2,3), (5,1), (4,2), (3,3), (5,2), (4,3) & (5,3)
= 0.23550	(1,0), (0,1), (2,0), (1,1), (0,2), (3,0), (2,1), (1,2), (0,3), (4,0), (3,1), (2,2), (1,3), (5,0), (4,1), (3,2), (2,3), (5,1), (4,2), (3,3), (5,2), (4,3) & (5,3)
= 0.23552	(0,0), (1,0), (0,1), (2,0), (1,1), (0,2), (3,0), (2,1), (1,2), (0,3), (4,0), (3,1), (2,2), (1,3), (5,0), (4,1), (3,2), (2,3), (5,1), (4,2), (3,3), (5,2), (4,3) & (5,3)

➤ Estimation Procedure:

A. Bayes Procedure:

Now the Bayesian approach is applied on this distribution. Using Bayesian method the Bayes estimate of the involved parameter, p is found. Then the Bayes risk and risk in case of original distribution are computed.

For this purpose, let us consider

$$\left. \begin{array}{l} X \sim B(n, p) \\ Y \sim B(m, p^2) \end{array} \right\} independently \ , 0 < p < 1$$

and let

$$p \sim Beta(\alpha, \beta)$$

$$\therefore \pi(p) = prior \ \ distribution \ \ of \ \ p$$

$$= \frac{1}{B(\alpha, \beta)} p^{\alpha-1}(1-p)^{\beta-1}$$

$\therefore$ The posterior distribution of p is

$$= q(p \mid x, y) = \frac{f(x, y \mid p)\pi(p)}{\int\limits_0^1 f(x, y \mid p)\pi(p)dp}$$

$$= \frac{p^{x+2y+\alpha-1}(1-p)^{n-x+m-y+\beta-1}(1+p)^{m-y}}{\int\limits_0^1 p^{x+2y+\alpha-1}(1-p)^{n-x+m-y+\beta-1}(1+p)^{m-y}dp} \qquad (3.1)$$

Now substituting the binomial expansion of $(1+p)^{m-y}$ in the integration portion of the expression (3.1), it can be written as

$$\int_0^1 p^{x+2y+\alpha-1}(1-p)^{n-x+m-y+\beta-1}(1+p)^{m-y}dp$$

$$=\int_0^1 p^{x+2y+\alpha-1}(1-p)^{n-x+m-y+\beta-1}[1+(m-y)p+{}^{(m-y)}c_2 p^2+{}^{(m-y)}c_3 p^3+...]dp$$

$$=\int_0^1 p^{x+2y+\alpha-1}(1-p)^{n-x+m-y+\beta-1}dp+(m-y)\int_0^1 p^{x+2y+\alpha+1-1}(1-p)^{n-x+m-y+\beta-1}dp+$$

$${}^{(m-y)}c_2\int_0^1 p^{x+2y+\alpha+2-1}(1-p)^{n-x+m-y+\beta-1}dp+{}^{(m-y)}c_3\int_0^1 p^{x+2y+\alpha+3-1}(1-p)^{n-x+m-y+\beta-1}dp+... \qquad (3.2)$$

From (3.1) and (3.2)

$$q(p\,|\,x,y)=\frac{p^{x+2y+\alpha-1}(1-p)^{n-x+m-y+\beta-1}(1+p)^{m-y}}{\begin{array}{l}B(x+2y+\alpha,n-x+m-y+\beta)+(m-y)B(x+2y+1+\alpha,n-x+m-y+\beta)+\\{}^{(m-y)}c_2 B(x+2y+2+\alpha,n-x+m-y+\beta)+{}^{(m-y)}c_3 B(x+2y+3+\alpha,n-x+m-\\ y+\beta)+...\end{array}}$$

$$=\frac{p^{x+2y+\alpha-1}(1-p)^{n-x+m-y+\beta-1}(1+p)^{m-y}}{L},\quad say$$

Using this posterior distribution the maximum likelihood estimate of p can be obtained.

The following steps will found the Bayesian estimation from the posterior distribution.

Under SEL, the Bayes estimation with respect to π(p) is

$$d_0(x,y)=E^{q(p|x,y)}(p)$$

$$=E(p\,|\,p\sim q(p\,|\,x,y)$$

$$=\int_0^1 pq(p\,|\,x,y)dp$$

$$=\frac{1}{L}\int_0^1 pp^{x+2y+\alpha-1}(1-p)^{n-x+m-y+\beta-1}(1+p)^{m-y}dp$$

$$=\frac{1}{L}\int_0^1 p^{x+2y+\alpha+1-1}(1-p)^{n-x+m-y+\beta-1}[1+(m-y)p+{}^{(m-y)}c_2 p^2+{}^{(m-y)}c_3 p^3+...]dp$$

$$\therefore E(p\,|\,p\sim q(p\,|\,x,y))$$

$$=\frac{1}{L}[B(x+2y+\alpha+1,n-x+m-y+\beta)+(m-y)B(x+2y+\alpha+2,n-x+m-y+\beta)$$

$$+{}^{(m-y)}c_2 B(x+2y+\alpha+3,n-x+m-y+\beta)+{}^{(m-y)}c_3 B(x+2y+\alpha+4,n-x+m-y+\beta)+...]$$

So, under square error loss the Bayes Estimate of p with respect to prior π(p) is

$$\frac{\begin{array}{l}B(x+2y+\alpha+1,n-x+m-y+\beta)+(m-y)B(x+2y+\alpha+2,n-x+m-y+\beta)\\+{}^{(m-y)}c_2 B(x+2y+\alpha+3,n-x+m-y+\beta)+{}^{(m-y)}c_3 B(x+2y+\alpha+4,n-x+m-y+\beta)+...\end{array}}{\begin{array}{l}B(x+2y+\alpha,n-x+m-y+\beta)+(m-y)B(x+2y+1+\alpha,n-x+m-y+\beta)+\\{}^{(m-y)}c_2 B(x+2y+2+\alpha,n-x+m-y+\beta)+{}^{(m-y)}c_3 B(x+2y+3+\alpha,n-x+m-\\ y+\beta)+...\end{array}}$$

$$\qquad (3.3)$$

= C (say).

B. *Maximum Likelihood Estimation:*

This estimation is done and also a comparison is given in the following table.

The maximum likelihood of p in classical method and Bayesian method and also the Bayes estimates of p with respect to the prior π(p) is given in the following table (taking x=0-5, y=0-3,α=2, β=3):

Values of (x, y)	Accepted values of p in classical method [ML estimates]	Accepted values of p in Bayesian method (i.e. in the posterior distribution) [ML estimates]	Baye's estimates of p
(0,0)	NA (0)	0.1159392	0.1796218
(0,1)	0.2559191	0.2761045	0.3071429
(0,2)	0.4171725	0.3977991	0.4107143
(0,3)	0.5454764	0.5	0.5000000
(1,0)	0.1702842	0.2201958	0.2629310
(1,1)	0.3711419	0.3615899	0.3798077
(1,2)	0.5164583	0.4745179	0.4772727
(1,3)	0.6363635	0.5714321	0.5625000
(2,0)	0.3112872	0.3175872	0.3433014
(2,1)	0.4818915	0.4451788	0.4513889
(2,2)	0.6146623	0.5507349	0.5434783
(2,3)	0.7272721	0.6428568	0.6250000
(3,0)	0.4391662	0.4104344	0.4213710
(3,1)	0.5893767	0.5272325	0.5220588
(3,2)	0.7119901	0.6264452	0.6093750
(3,3)	0.8181791	0.7142792	0.6875000
(4,0)	0.5592660	0.5000000	0.4975962
(4,1)	0.6944718	0.6080283	0.5919540
(4,2)	0.8085302	0.7017674	0.6750000
(4,3)	0.9090926	0.785716	0.7500000
(5,0)	0.6742002	0.5870874	0.5723140
(5,1)	0.7977114	0.6878677	0.6611842
(5,2)	0.9045345	0.7767435	0.7403846
(5,3)	0.0000000	0.8571339	0.8125000

$$\therefore V^{q(p|x,y)}(p)$$

$$= E^{q(p|x,y)}(p-C)^2$$

$$= \int_0^1 (p-C)^2 q(p|x,y)dp, \quad where \ C \ is \ a \ p \ free \ quantity$$

$$= \frac{1}{L}\int_0^1 (p^2 - 2pC + C^2) p^{x+2y+\alpha-1}(1-p)^{n-x+m-y+\beta-1}(1+p)^{m-y} dp$$

$$= \frac{1}{L}[\int_0^1 p^{x+2y+\alpha+2-1}(1-p)^{n-x+m-y+\beta-1}(1+p)^{m-y} dp - 2C\int_0^1 p^{x+2y+\alpha+1-1}(1-p)^{n-x+m-y+\beta-1}(1+p)^{m-y} dp + C^2$$

$$\int_0^1 p^{x+2y+\alpha-1}(1-p)^{n-x+m-y+\beta-1}(1+p)^{m-y} dp]$$

$$= \frac{1}{L}[B(x+2y+2+\alpha, n-x+m-y+\beta) + (m-y)B(x+2y+3+\alpha, n-x+m-y+\beta) +$$

$$^{(m-y)}c_2 B(x+2y+4+\alpha, n-x+m-y+\beta) + {}^{(m-y)}c_3 B(x+2y+5+\alpha, n-x+m-y+\beta) + ...$$

$$-2C\{B(x+2y+1+\alpha, n-x+m-y+\beta) + (m-y)B(x+2y+2+\alpha, n-x+m-y+\beta) +$$

$$+{}^{(m-y)}c_2 B(x+2y+3+\alpha, n-x+m-y+\beta) + {}^{(m-y)}c_3 B(x+2y+4+\alpha, n-x+m-y+\beta) + ...\}$$

$$+C^2\{B(x+2y+\alpha, n-x+m-y+\beta) + (m-y)B(x+2y+1+\alpha, n-x+m-y+\beta) + {}^{(m-y)}c_2$$

$$B(x+2y+2+\alpha, n-x+m-y+\beta) + {}^{(m-y)}c_3 B(x+2y+3+\alpha, n-x+m-y+\beta) + ...\}]$$

$$= \frac{V}{L} \ (say) \tag{3.4}$$

Hence the Bayes risk is given by

$$\iint_{x,y} \binom{n}{x}\binom{m}{y} L \frac{V}{L} dx dy$$

$$= \iint_{x,y} \binom{n}{x}\binom{m}{y} V dx dy$$

=0.0007318181 with absolute error < 1e-04   (using R software)

Now, the risk in case of the joint distribution of X and Y are to be found. For this purpose $E^{f(x,y|p)}(p- E^{f(x,y|p)}(p))^2$ is to be computed.

$E^{f(x,y|p)}(p-E^{f(x,y|p)}(p))^2 = E(p-d_i)^2$ where d_i is the likelihood estimate of p for the ith combination of X and Y.

Hence the risk in case of joint distribution of X and Y can be found in the following way.

$$E(p-d_i)^2$$

$$= \sum_{x,y} \int_0^1 (p-d_i)^2 f(x,y \mid p) \pi(p) dp$$

$$= \sum_{x,y} \binom{n}{x}\binom{m}{y} \int_0^1 (p-d_i)^2 p^{x+2y+\alpha-1}(1-p)^{n-x+m-y+\beta-1}(1+p)^{m-y} dp$$

The integration can be solved as the earlier computations [e.g. the procedure followed to find (3.4)] using various values of d_i's.
Finally, using R software

$$\sum_{x,y} \int_0^1 (p-d_i)^2 f(x,y \mid p) \pi(p) dp = 0.002524987$$

4. *On Bivariate Normal:-*

Let us suppose $X = (X_1, X_2)$ has a Bivariate normal distribution with zero mean, standard deviation equal to 1 and correlation parameter ρ.
Then the density of X is

$$f(x \mid \rho) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2(1-\rho^2)}\left(x_1^2+x_2^2 - 2\rho x_1 x_2\right)\right] I x_1, x_2 \in R$$

$$= \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left[-\frac{x_1^2+x_2^2}{2(1-\rho^2)} + \frac{\rho}{1-\rho^2}x_1 x_2\right] I x_1, x_2 \in R \tag{4.1}$$

Therefore

$$\eta_1 = -\frac{1}{2(1-\rho^2)} \qquad T_1(x) = x_1^2 + x_2^2$$

$$\eta_2 = \frac{\rho}{1-\rho^2} \qquad T_2(x) = x_1 x_2$$

$$\Sigma_\theta = \begin{pmatrix} V(T_1) & \mathrm{cov}(T_1,T_2) \\ \mathrm{cov}(T_1,T_2) & V(T_2) \end{pmatrix}$$

To compute the statistical curvature of this distribution firstly the moment generating function of bivariate normal distribution and its corresponding calculations are considered.
The moment generating function of $X \sim N(0, 0, 1, 1, \rho)$ is given by

$$M(t_1, t_2) = \exp\left[\frac{1}{2}(t_1^2 + 2\rho t_1 t_2 + t_2^2)\right]$$

$$\Leftrightarrow \sum_{i=0}^\infty \sum_{j=0}^\infty \mu_{ij} \frac{t_1^i t_2^j}{i!j!} = \sum_{k=0}^\infty \frac{(t_1^2 + 2\rho t_1 t_2 + t_2^2)^k}{2^k k!} \tag{4.2}$$

Equating the coefficients from both sides of (4.2), the bivariate moments are

$\mu_{11} =$ coefficient of $t_1 t_2 = \rho$

$\mu_{20} =$ coefficients of $t_1^2/2! = 1$

$\mu_{02} =$ coefficient of $\dfrac{t_2^2}{2!} = 1$

$\mu_{40} =$ coefficient of $\dfrac{t_1^4}{4!} = 3$

$\mu_{04} =$ coefficient of $\dfrac{t_2^4}{4!} = 3$

$\mu_{22} =$ coefficient of $\dfrac{t_1^2}{2!}\dfrac{t_2^2}{2!} = 2\rho^2 + 1$

Also, $\mu_{ij} = 0$ if $i+j$ is odd.
Now $V(T_1) = V(x_1^2 + x_2^2)$

$$= V(x_1^2) + V(x_2^2) + 2\mathrm{cov}(x_1^2, x_2^2)$$

Now, $V(x_1^2) = E(x_1^4) - E^2(x_1^2) = \mu_{40} - \mu_{20}^2 = 3 - 1 = 2$

$V(x_2^2) = 2$

&

$\mathrm{cov}(x_1^2, x_2^2) = E(x_1^2 x_2^2) - E(x_1^2)E(x_2^2) = \mu_{22} - \mu_{20}\mu_{02} = 2\rho^2$

$$\therefore V(T_1) = 2 + 2 + 4\rho^2 = 4(1 + \rho^2) \quad (4.3)$$

$$V(T_2) = V(x_1 x_2) = E(x_1^2 x_2^2) - E^2(x_1 x_2) = 1 + \rho^2 \quad (4.4)$$
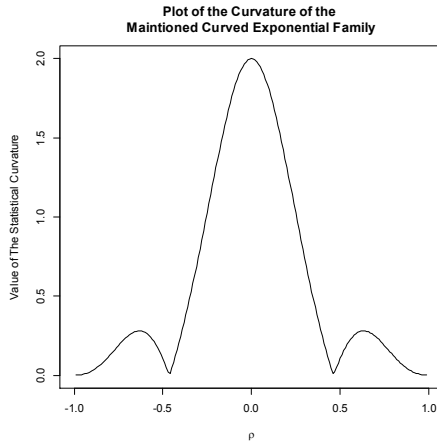
And after calculation,

$$\mathrm{Cov}(T_1, T_2) = 4\rho \quad (4.5)$$

$$\therefore \Sigma_\theta = \begin{pmatrix} 4(1 + \rho^2) & 4\rho \\ 4\rho & 1 + \rho^2 \end{pmatrix} \quad (4.6)$$

With this information an R program is used to find the values of the curvature.

The curvature of the mentioned distribution is given in the following figure:

Figure3

Plot of the Curvature of the Maintioned Curved Exponential Family

From the above figure it is observed that when the values of ρ lie between (-1,-0.5) and (0.5, 1) then this curved family performs just like the corresponding exponential family. But there is a good test statistic for bivariate normal distribution, belongs to exponential family, when the value of ρ_0 (the value for null hypothesis) is equal to 0. So, in this case a good statistic can not be found.

Remark:

The paper hopes, this can be done in other discrete curved family also and plenty of scope are there for doing further research.

References

1. Defining the Curvature of a Statistical Problem (With Applications to Second Order Efficiency). Bradley Efron. The Annals of Statistics, Vol. 3, No. 6. (Nov., 1975), pp, 1189-1242

2. Paremeterizing And Scoring Mixed Ancestral Graphs. Thomas Richardson & Peter Spirtes. Technical Report No. CMU-PHIL-102, August 1999.

3. Curved Exponential Family Models for Social Networks. David. R. Hunter, Pann State University, March 2, 2006.

4. Notes on Exponential Family. A. Das Gupta, Professor of Statistics, Purdue University. (http://www.stat.purdue.edu/~dasgupta/expfamily.pdf)

5. Lazega's Lawyer Dataset. Lazega & Pattison (1999) and Lazega (2001).

6. Testing Statistical Hypotheses, Third Edition. E. L. Lehmann, Joseph P. Romano.

7. Statistical Software Used:- R