The Korean Communications in Statistics **Vol. 14** No. 1, 2007, pp. 147–153

A Note Based on Multiparameter Discrete Exponential Families in View of Cacoullos-type Inequalities*

G. R. Mohtashami Borzadaran[1)]

Abstract

In this note, we obtained results related to multiparameter discrete exponential families on considering lattice or semi-lattice in place of N (Natural numbers) in view of Cacoullos-type inequalities via the same arguments in Papathanasiou (1990, 1993).

Keywords: Chernoff-type Inequalities; variance bounds; characterization; upper bounds; exponential families; lower bounds.

1. Introduction

There is an extensive literature dealing with upper and lower bounds for the variance of function of a random variable. The starting point dates back to the result of Chernoff (1981), which gives a bound for the variance of an absolutely continuous function (w.r.t. Lebesgue measure) of a normal random variable. Chen (1982), Cacoullos (1982, 1989), Klaassen (1985) and Borovkov and Utev (1984) obtained characterizations related to Chernoff-type inequalities. During the last fifteen years or so, several papers have appeared on modified versions or variants of the Chernoff inequality, and related characterizations. Variations or extended versions of these latter results have been obtained by Cacoullos and Papathanasiou (1985, 1986, 1989, 1995, 1997), Koicheva (1993), Alharbi and Shanbhag (1996), Mohtashami Borzadaran and Shanbhag (1998) and others. Papathanasiou (1990, 1993) characterized a version of discrete exponential family via Chernoff-type inequalities. We extend the idea of Papathanasiou (1990, 1993) based on considering lattice or semi-lattice in place of N "Natural numbers".

* This research is supported by the Statistics Excellence Center of University of Mashhad, Mashhad-Iran.

1) Assistant Professor, Department of Mathematics and Statistics Faculty of Science, University of Birjand, Birjand, Iran.
E-mail : Gmohtashami@birjand.ac.ir

2. Main Result

Let the distribution of a random vector $\underline{X}$ belongs to a multidimensional family. Under some assumptions, a Cacoullos-type inequality for the variance of $g(\underline{X}), \underline{X} \in N^p$ considered by Papathanasiou (1990, 1993) for characterizing an exponential family. Consider now the multiparameter discrete exponential family:

$$f(\underline{x}) \propto \theta_1^{\frac{x_1}{\beta}} . \theta_2^{\frac{x_2}{\beta}} \cdots \theta_p^{\frac{x_p}{\beta}} k(\underline{x}), \ \underline{x} = (x_1, \ldots, x_p) \in B^{*p}, \tag{2.1}$$

such that $\theta_i > 0$, $x \in B^* = \{x : x = n\beta + \alpha, n \in \mathcal{N}\}$ with $\beta > 0, \alpha \in R$, where $\mathcal{N}$ is N (natural numbers) or Z (integer numbers) or $\{0, 1, 2, \ldots, n_0\}$. Also, let g be a real function defined on B^{*p} as:

$$t_i^*(\underline{x}) = \frac{k(x_1, x_2, \ldots, x_{i-1}, x_i - \beta, x_{i+1}, \ldots, x_p)}{k(\underline{x})}, \tag{2.2}$$

where $k(.) > 0$ is a real-valued function and:

$$\Delta_i^\beta g(\underline{x}) = \frac{g(x_1, x_2, \ldots, x_{i-1}, x_i + \beta, x_{i+1}, \ldots, x_p) - g(\underline{x})}{\beta}, \ \beta > 0. \tag{2.3}$$

Suppose that

$$E||\Delta_1^\beta g, \Delta_2^\beta g, \ldots, \Delta_p^\beta|| < \infty \tag{2.4}$$

and

$$E||t_1^* g, t_2^* g, \ldots, t_p^* g|| < \infty, \tag{2.5}$$

where $||.||$ means the norm of the vector.

Note that if x_i is the infimum of all points B^{*p} , then $k(x_1, x_2, \ldots, x_{i-1}, x_i - \beta, x_{i+1}, \ldots, x_p) = 0$ and if x_i is the supremum of all points in B^{*p} then $g(x_1, x_2, \ldots, x_{i-1}, x_i + \beta, x_{i+1}, \ldots, x_p) = g(\underline{x})$. The following assertions can be obtained via the same arguments in Papathanasiou (1993) on considering lattice or semi-lattice in place of N:

Theorem 2.1 *Under the preceding conditions, let $g(.)$ be as above, then*

$$\operatorname{Cov}\{t_i^*(\underline{X}), g(\underline{X})\} = \beta\theta_i E\{\Delta_i^\beta g(\underline{X})\}, \underline{X} \in B^{*p}, i = 1, 2, \ldots, p. \tag{2.6}$$

Proof: Via the notation similar to those in Papathanasiou (1990), we have,

$$\begin{aligned}
\mathrm{Cov}\{t_i^*(\underline{X}), g(\underline{X})) &= \sum_{x_1}\cdots\sum_{x_p} t_i^*(\underline{x})g(\underline{x})f(\underline{x}) - E(t_i^*(\underline{X}))E(g(\underline{X})) \\
&= \theta_i \sum_{y_1}\cdots\sum_{y_i}\cdots\sum_{y_p} k(y_1,\ldots,y_p)c(\theta_1,\ldots,\theta_p)\theta_1^{\frac{y_1}{\beta}}.\theta_2^{\frac{y_2}{\beta}}\cdots\theta_p^{\frac{y_p}{\beta}} \\
&\quad \times g(y_1,\ldots,y_{i-1},y_i+\beta,y_{i+1},\ldots,y_p) - \theta_i E(g(\underline{X})) \\
&= \theta_i[\beta E(\Delta_i^\beta g(\underline{X})) + E(g(\underline{X}))] - \theta_i E(g(\underline{X})) \\
&= \theta_i \beta E(\Delta_i^\beta g(\underline{X})),
\end{aligned}$$

where $x_j = y_j$ for $j \neq i$ and $x_j = y_j + \beta$ for $j = i$ and

$$c(\theta_1,\ldots,\theta_p) = \left[\sum_{x_1}\cdots\sum_{x_p}\theta_1^{\frac{x_1}{\beta}}.\theta_2^{\frac{x_2}{\beta}}\cdots\theta_p^{\frac{x_p}{\beta}}k(\underline{x})\right]^{-1}$$

is the normalizing constant. □

Theorem 2.2 *We have $I_{ij}^* = \beta\theta_j^{-1}E\{\Delta_i^\beta t_j^*(\underline{X})\}, \underline{X} \in B^{*p}$, $i,j = 1,2,\ldots,p$, such that I_{ij}^* is the $(i,j)^{th}$ element of the $p \times p$ analogue of Fisher information matrix $I_{\underline{X}}^* = (I_{ij}^*), i,j = 1,2,\ldots,p$, where*

$$I_{ij}^* = E\left\{\frac{f(\underline{X}) - f_i(\underline{X})}{f(\underline{X})}\right\}\left\{\frac{f(\underline{X}) - f_j(\underline{X})}{f(\underline{X})}\right\}, \underline{X} \in B^{*p}$$

and $f_i(\underline{x}) = f(x_1, x_2, \ldots, x_{i-1}, x_i - \beta, x_{i+1}, \ldots, x_p)$ for $i,j = 1,2,\ldots,p$.

Proof: Using theorem 2.1 implies that,

$$\begin{aligned}
E(t_i^*(\underline{X})t_j^*(\underline{X})) &= E(t_i^*(\underline{X}))E(t_j^*(\underline{X})) + \mathrm{Cov}(t_i^*(\underline{X}), t_j^*(\underline{X})) \\
&= \theta_i\theta_j + \beta\theta_i E(\Delta_i^\beta t_j^*(\underline{X})).
\end{aligned}$$

We have,

$$\begin{aligned}
I_{ij}^* &= E\left\{\frac{f(\underline{X}) - f_j(\underline{X})}{f(\underline{X})}\right\}\left\{\frac{f(\underline{X}) - f_i(\underline{X})}{f(\underline{X})}\right\} \\
&= 1 - E\left(\frac{f_j(\underline{X})}{f(\underline{X})}\right) - E\left(\frac{f_i(\underline{X})}{f(\underline{X})}\right) + E\left(\frac{f_i(\underline{X})}{f(\underline{X})}, \frac{f_j(\underline{X})}{f(\underline{X})}\right).
\end{aligned}$$

On noting that

$$E\left(\frac{f_i(\underline{X})}{f(\underline{X})}\right) = \sum_{x_1}\cdots\sum_{x_p} f(x_1,\ldots,x_{i-1},x_i-\beta,x_{i+1},\ldots,x_p)$$
$$= \frac{1}{\theta_i}\sum_{x_1}\cdots\sum_{x_p}\frac{h(x_1,\ldots,x_{i-1},x_i-\beta,x_{i+1},\ldots,x_p)}{h(x_1,\ldots,x_p)}f(x_1,\ldots,x_p)$$
$$= \frac{1}{\theta_i}E(t_i^*(\underline{X})),$$

and

$$E\left(\frac{f_i(\underline{X})}{f(\underline{X})}\cdot\frac{f_j(\underline{X})}{f(\underline{X})}\right) = \frac{1}{\theta_i\theta_j}\sum_{x_1}\cdots\sum_{x_p}\frac{h(x_1,\ldots,x_{i-1},x_i-\beta,x_{i+1},\ldots,x_p)}{h(x_1,\ldots,x_p)}$$
$$\times\frac{h(x_1,\ldots,x_{j-1},x_j-\beta,x_{j+1},\ldots,x_p)}{h(x_1,\ldots,x_p)}f(x_1,\ldots,x_p)$$
$$= \frac{1}{\theta_i\theta_j}E(t_i^*(\underline{X}).t_j^*(\underline{X})).$$

We obtain that,

$$I_{ij}^* = \beta\theta_j^{-1}E\{\Delta_i^\beta t_j^*(\underline{X})\}.$$

□

Theorem 2.3 (Papathanasiou (1990)): *Let $\underline{X} = (X_1,\ldots,X_p)$ have continuous p-dimensional family with density $f(x) = c(\theta)e^{\theta x - K(x)}$ where $\theta = (\theta_1,\ldots,\theta_p)$ and g be any function such that*

$$E(|\nabla g(\underline{X})|) < \infty,\ E\{|(\nabla k(\underline{X}) - \theta)g(\underline{X})|\} < \infty,$$

and

$$E\{|(\nabla k(\underline{X}) - \theta)g(\underline{X})|\} = E(\nabla g(\underline{X})),$$

then, we have,

$$V(g(\underline{X})) \geq E(\nabla g(\underline{X}))I^{*-1}E(\nabla g(\underline{X})), \tag{2.7}$$

such that I^{*-1} is the inverse of

$$I^* = (K_{ij}),\ K_{ij} = E[k_{ij}(X)],\ k_{ij} = \frac{\partial K(\underline{x})}{\partial x_i \partial x_j},$$

and I^* is assumed to be non-singular. Equality holds in (2.7) iff g is linear in

$$k_i(x) = D_i K(x) = \frac{\partial K(x)}{\partial x_i},\ i = 1,\ldots,p.$$

Proof: We have the multivariate Cauchy-Schwartz inequality (Cacoullos, 1989),

$$V(g(\underline{X})) \geq \text{Cov}(g(\underline{X}), \nabla k(\underline{X}))[D(\nabla k(\underline{X}))]^{-1}\text{Cov}(g(\underline{X}), \nabla k(\underline{X}))$$

and

$$\begin{aligned}\text{Cov}(g(\underline{X}), k_i(\underline{X})) &= E(g_i(\underline{X})),\\ \text{Cov}(k_i(\underline{X}), k_j(\underline{X})) &= E(k_{ij}(\underline{X})).\end{aligned}$$

Hence, the proof is complete. □

Theorem 2.4 *Under the preceding conditions (2.1),(2.2),(2.3),(2.4),(2.5), let g be defined (arbitrary real- valued function) as above, then*

$$\begin{aligned}\text{Var}[g(\underline{X})] \geq\ & E_\theta\{\theta_1\Delta_1^\beta g, \theta_2\Delta_2^\beta g, \ldots, \theta_p\Delta_p^\beta g\}^t(I_{\underline{X}}^*)^{-1}\\ &\times E_\theta\{\theta_1\Delta_1^\beta g, \theta_2\Delta_2^\beta g, \ldots, \theta_p\Delta_p^\beta g\},\ \underline{X} \in B^{*p} \qquad (2.8)\end{aligned}$$

with equality in (2.8) if and only if g is linear in $t_j^*(.)$, where $I_{\underline{X}}^*$ is the analogue of the Fisher information matrix.

Proof: It is easy to follow the arguments in Papathanasiou (1990) as special case of the above theorem with lattice case in place of the Natural numbers. □

Theorem 2.5 *Let inequality (2.8) is satisfied for every real-valued function g defined on B^{*p} with equality for g linear in $t_j^*(\underline{x})$ where $t_j^*(\underline{x})$ is given by (2.2) with $E\{t_j^*(\underline{X})\} = \theta_j$, $j = 1, 2, \ldots, p$ and (2.4) and (2.5) are satisfied. Then the probability distribution of $\underline{X}$ is given by (2.1).*

Proof: Following the proof of the theorem 2.2 in Papathanasiou (1990) by setting $g(\underline{x}) = t_j(\underline{x}) + \lambda q(\underline{x})$, we get the identity (2.6) for $q(\underline{x}) = s_1^{x_1/\beta} \cdots s_p^{x_p/\beta}$ and obtain the desired characterizations. □

Remark 2.1 The proof of the above assertions can be obtained via the same argument in Papathanasiou (1990, 1993) with lattice case in place of the Natural numbers. We can have Papathanasiou's results on taking $\beta = 1$ and $n \in N$.

Remark 2.2 In (2.1), if $h(\underline{x}) = \prod_{i=1}^n h_i^*(x_i), h_i^*(x_i) > 0$, then $f(\underline{x})$ is the joint probability density function of n independent random variable with $t_j^*(x_j) = \{h_j^*(x_j - \beta)\}/\{h_j^*(x_j)\}$, $x_j \in B^*$. Here is an example of characterizations of the bivariate bilateral power series via the above assertions that showed in Table 2.1:

Table 2.1: Characterizations of the Bivariate Bilateral Polynomial Power Series families

$t_i^*(x_i)\downarrow\ t_j^*(x_j)\rightarrow$	$q^{-(\frac{x_j}{\beta}-1)(\frac{x_j}{\beta}-2)\cdots(\frac{x_j}{\beta}-k+1)}$	$q^{-(\frac{x_j}{\beta}-1)(\frac{x_j}{\beta}-2)(\frac{x_j}{\beta}-3)}$	$q^{-(\frac{x_j}{\beta}-1)}$
$q^{-(\frac{x_i}{\beta}-1)(\frac{x_i}{\beta}-2)\cdots(\frac{x_i}{\beta}-k+1)}$	BPPD-BPPD	BPPD-DQ	BPPD-DN
$q^{-(\frac{x_i}{\beta}-1)(\frac{x_i}{\beta}-2)(\frac{x_i}{\beta}-3)}$	DQ-BPPD	DQ-DQ	DQ-DN
$q^{-(\frac{x_i}{\beta}-1)}$	DN-BPPD	DN-DQ	DN-DN

Note that in the above table BPPD, DQ and DN denote the bilateral polynomial power series, discrete quartic and discrete normal distributions respectively, and symbols such as DQ-DN are used to denote the bivariate discrete quartic-discrete normal distributions; DQ-BPPD here stand for the bivariate discrete quartic-bilateral power series distribution and so on. For details of them, see Kemp (1997) and Mohtashami Borzadaran (2000).

3. Conclusions

In this note in view of Papthanasiou (1990, 1993), extended some ideas to lattice or semi-lattice cases. Also obtained an example in lattice and semi-lattice cases based on Papthanasiou (1990, 1993).

Acknowledgments

I am grateful to Prof. D. N. Shanbhag for his evaluable suggestion.

References

Alharbi, A. A. and Shanbhag, D. N. (1996). General characterization theorems based on versions of the Chernoff inequality and the Cox representation. *Journal of Statistical Planning and Inference*, **55**, 139–150.

Borovkov, A. A. and Utev, S. A. (1984). On an inequality and a related characterization of the normal distribution. *Theory of Probability and its Applications*, **28**, 219–228.

Cacoullos, T. (1982). On upper and lower bounds for the variance of a function of a random variable. *Annals of Probability*, **10**,799–809.

Cacoullos, T. (1989). Dual Poincaré-type inequalities via the Cramér-Rao and the Cauchy-Schwarz inequalities and related characterizations. *Statistical Data Analysis and Inference*, Nortg-Holland, Amsterdam, 239-250.

Cacoullos, T. and Papathanasiou, V. (1985). On upper bounds for the variance of functions of random variables. *Statistics & Probability Letters*, **3**, 175–184.

Cacoullos, T. and Papathanasiou, V. (1986). Bounds for the variance of functions of random variables by orthogonal polynomial and Bhattacharyya bounds. *Statistics & Probability*

Letters, **4**, 21-23.

Cacoullos, T. and Papathanasiou, V. (1989). Characterizations of distributions by variance bounds. *Statistics & Probability Letters*, **7**, 351–356.

Cacoullos, T. and Papathanasiou, V. (1995). A generalization of covariance identity and related characterizations. *Mathematical Methods of Statistics*, **4**, 106–113.

Cacoullos, T. and Papathanasiou, V. (1997). Characterizations of distributions by generalizations of variance bounds and simple proofs of the CLT. *Journal of Statistical Planning and Inference*, **63**, 157–171.

Chen, L. H. Y. (1982). An inequality for the multivariate normal distribution. *Journal of Multivariate Analysis*, **12**, 306–315.

Chernoff, H. (1981). A note on an inequality involving the normal distribution. *The Annals of Probability*, **9**, 533–535.

Kemp, A. W. (1997). Characterizations of a discrete normal distribution. *Journal of Statistical Planning and Inference*, **63**, 223–229.

Klaassen, C. A. J. (1985). On an inequality of Chernoff. *The Annals of Probability*, **13**, 966–974.

Koicheva, M. (1993). On an inequality and the related characterizations of the gamma distribution. *Applications of Mathematics*, **38**, 11–18.

Mohtashami Borzadaran, G. R. (2001). Characterization of Pearsonian and bilateral power series distribution via maximum entropies. *Bayesian inference and maximum entropy methods in science and engineering, AIP Conference Proceedings*, **568**, 145–150.

Mohtashami Borzadaran, G. R. and Shanbhag, D. N. (1998). Further results based on Chernoff-type inequalities. *Statistics & Probability Letters*, **39**, 109–117.

Papathanasiou, V. (1990). Characterizations of multidimensional exponential families by Cacoullos-type inequalities. *Journal of Multivariate Analysis*, **35**, 102–107.

Papathanasiou, V. (1993). Some characteristic properties of the Fisher information matrix via Cacoullos-type inequalities. *Journal of Multivariate Analysis*, **44**, 256–265.

[Received February 2006, Accepted January 2007]