行政院國家科學委員會補助專題研究計畫 ■成果報告 □期中進度報告

Dirichlet隨機向量線性組合及其貝氏應用的研究

A study on the linear combination of Dirichlet vector and its Bayesian applications

計畫類別：■個別型計畫　□整合型計畫
計畫編號：NSC 91-2118-M-004-001
執行期間：91年8月1日至92年7月31日

計畫主持人：姜志銘
計畫參與人員：郭錕霖　柯力文　陳力揚　劉桂芳

成果報告類型(依經費核定清單規定繳交)：■精簡報告　□完整報告

本成果報告包括以下應繳交之附件：
□赴國外出差或研習心得報告一份
□赴大陸地區出差或研習心得報告一份
■出席國際學術會議心得報告及發表之論文各一份
□國際合作研究計畫國外研究報告書一份

處理方式：除產學合作研究計畫、提升產業技術及人才培育研究計畫、列管計畫及下列情形者外，得立即公開查詢
□涉及專利或其他智慧財產權，□一年□二年後可公開查詢

執行單位：國立政治大學應用數學系

中華民國九十二年十月十三日

行政院國家科學委員會專題研究計畫成果報告

Dirichlet 隨機向量線性組合及其貝氏應用的研究
A study on the linear combination of Dirichlet vector and its Bayesian applications

計畫編號：NSC 91-2118-M-004-001
執行期限：91年8月1日至92年7月31日
主持人：姜志銘　國立政治大學應用數學系
jiangt@math.nccu.edu.tw
計畫參與人員：郭錕霖　柯力文　陳力揚　劉桂芳
國立政治大學應用數學系

中文摘要

直方類別的未知抽樣機率的貝氏區域平滑性在很多領域的應用上是很重要的。Dickey and Jiang (1998) 首先提出過濾變量的Dirichlet分配作為貝氏區域平滑性的應用。我們提供新的工具，探討，並得到Ferguson-Dirichlet過程的隨機函數的新結果；過濾變量的Dirichlet分配可算是它的特殊情形。

關鍵詞：貝氏推論，粗糙性，過濾變量Dirichlet分配，Ferguson-Dirichlet過程

Abstract

Bayesian local smoothness of the unknown sampling probabilities of histogram categories is important for a wide range of applications. Dickey and Jiang (1998) first introduced the filtered-variate Dirichlet distribution for applications on Bayesian local smoothness. We provide new tools, study, and give the results of the random functional of a Ferguson-Dirichlet process, a special case of which is the filtered-variate Dirichlet distribution.

Keywords: Bayesian inference, roughness, filtered-variate Dirichlet distribution, Ferguson-Dirichlet process

1 Introduction

Bayesian local smoothness of the unknown sampling probabilities of histogram categories is important for a wide range of applications, from uses of one-way histogram data to medical diagnosis, optical image processing, and other uses of multidimensional histograms. This problem has been important for decades, see Dickey and Jiang (1998) for detailed discussion and references therein. Promising methods of prior assessment, based on filtered-variate Dirichlet distributions, were also given by Dickey and Jiang (1998). The one-dimensional distribution of the filtered-variate Dirichlet distribution has been studied by Hannum, Hollander, and Langberg (1981), Yamato (1984), Jiang (1984,1988,1991), Cifarelli and Regazzini (1990), and Diaconis and Kemperman (1996), among others.

In this research, we shall study a new one-dimensional distribution of the filtered-variate Dirichlet distribution, which shall be important by itself and for extention to spherical higher dimensional cases. We shall also study the higher dimensional distribution of the filtered-variate Dirichlet distribution.

2 Notations

Let $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_m)'$ be a Dirichlet vector with parameter vector $\boldsymbol{a} = (a_1, \ldots, a_m)'$, denoted by $\boldsymbol{\theta} \sim D(\boldsymbol{a})$, if the probability density function of $\boldsymbol{\theta}$ is

$$f(\boldsymbol{\theta}) = \frac{1}{B(\boldsymbol{a})} \prod_{i=1}^{m} \theta_i^{a_i - 1},$$

where $B(\boldsymbol{a}) = (\prod_{i=1}^{m} \Gamma(a_i))/\Gamma(a_+)$, and $a_+ = \sum_{i=1}^{m} a_i$, for all $\boldsymbol{\theta}$ in the probability simplex $\{\, \boldsymbol{\theta} \mid \text{each } \theta_i > 0,\ \theta_+ = 1 \,\}$. Define $\boldsymbol{\alpha} = G\boldsymbol{\theta}$ as a filtered-variate Dirichlet random vector, where G is a $k \times m$ constant matrix. Then a one-dimensional marginal random variable of $\boldsymbol{\alpha}$ is a linear combination of the coordinates of the random vector $\boldsymbol{\theta}$. For example, $\alpha_i = g_{i1}\theta_1 + g_{i2}\theta_2 + \cdots + g_{im}\theta_m$.

Let U be a Ferguson's (1973) Dirichlet process with parameter μ on (Θ, B), where $\Theta = [0, 2\pi)$, B is the σ-field of Borel subsets of Θ, and μ is a non-null finite measure on (Θ, B). We are interested in the random functional of $\int_\Theta g(\theta)\, U(d\theta)$, where $g(\theta) = (g_1(\theta), \ldots, g_k(\theta))$ is a k-variate function. Note that a filtered-variate Dirichlet distribution is a special case of $\int_\Theta g(\theta)\, U(d\theta)$. Here, we shall study the distribution of $\int_\Theta g(\theta)\, U(d\theta)$, in particular when $k = 2$ and μ is a uniform measure with $\mu(\Theta) = c$.

3 The c-characteristic functions and their properties

First, we define a new multivariate characteristic function called a multivariate c-characteristic function.

Definition 3.1 *If $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_m)'$ is a random vector on a subset S of $B = [-b_1, b_1] \times [-b_2, b_2] \times \cdots \times [-b_m, b_m]$, it's multivariate c-characteristic function is defined as*

$$g(\boldsymbol{t}; \boldsymbol{\theta}, c) = \mathop{E}_{\boldsymbol{\theta}}\left[(1 - i\boldsymbol{t} \cdot \boldsymbol{\theta})^{-c}\right], \quad |\,\boldsymbol{t}\,| < b^{-1}, \tag{3.1}$$

where c is a positive real number, $b = \sqrt{\sum_{i=1}^{m} b_i^2}$, $\boldsymbol{t}' = (t_1, \ldots, t_m)$, $|\,\boldsymbol{t}\,| = \sqrt{\sum_{i=1}^{m} t_i^2}$, and $\boldsymbol{t} \cdot \boldsymbol{\theta} = \sum_{i=1}^{m} t_i \theta_i$.

Note that the univariate c-characteristic function given by Jiang (1988, 1991) is a special case of this multivariate c-characteristic function. Next, we shall give some of its properties without proof. For more complete properties and their proofs, see Jiang, Dickey, and Kuo (2003). First, it has a one-to-one correspondence between $g(\boldsymbol{t}; \boldsymbol{\theta}, c)$ and $\boldsymbol{\theta}$.

Lemma 3.2 *For random variables $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$ with supports in a subset S of $B = [-b_1, b_1] \times [-b_2, b_2] \times \cdots \times [-b_m, b_m]$ and any positive real number c, if we have*

$$g(\boldsymbol{t}; \boldsymbol{\theta}_1, c) = g(\boldsymbol{t}; \boldsymbol{\theta}_2, c)$$

for all $|\,\boldsymbol{t}\,| < b^{-1}$, where $b = \sqrt{\sum_{i=1}^{m} b_i^2}$. Then

$$\boldsymbol{\theta}_1 \sim \boldsymbol{\theta}_2.$$

Lemma 3.3 *If $\boldsymbol{\theta}$ and the $\boldsymbol{\theta}_n$'s are random vectors with supports in S, their corresponding multivariate c-characteristic functions are $g(\boldsymbol{t}; \boldsymbol{\theta}, c)$ and $g(\boldsymbol{t}; \boldsymbol{\theta}_n, c)$, and the sequence of random vectors $\{\boldsymbol{\theta}_n\}$ converges in distribution to $\boldsymbol{\theta}$, then, for all $|\,\boldsymbol{t}\,| < b^{-1}$,*

$$g(\boldsymbol{t}; \boldsymbol{\theta}_n, c) \to g(\boldsymbol{t}; \boldsymbol{\theta}, c), \quad \text{as } n \to \infty.$$

Finally, we give the following important convergence theorem.

Theorem 3.4 *Assume $\boldsymbol{\theta}$ and $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \ldots$ are random vectors with supports in S and their corresponding multivariate c-characteristic function are $g(\boldsymbol{t}; \boldsymbol{\theta}, c), g(\boldsymbol{t}; \boldsymbol{\theta}_2, c), g(\boldsymbol{t}; \boldsymbol{\theta}_2, c), \ldots$, respectively. If, for all $|\,\boldsymbol{t}\,| < b^{-1}$,*

$$g(\boldsymbol{t}; \boldsymbol{\theta}_n, c) \to g(\boldsymbol{t}; \boldsymbol{\theta}, c), \quad \text{as } n \to \infty,$$

then the sequence of the random vector $\boldsymbol{\theta}_n$ converges in distribution to $\boldsymbol{\theta}$.

4 Distribution of a random functional of a Ferguson process over the region bounded by an ellipse

Let $\Theta = [0, 2\pi)$, a and b be positive real numbers, U be a Ferguson's process with parameter μ on (Θ, B), B be the σ-field of Borel subsets of Θ, and μ be a non-null (positive) finite measure on (Θ, B). Let $\beta(\theta) = (2\pi/c)\mu[0, \theta)$ with $c = \mu(\Theta)$. Then define a random vector $\boldsymbol{v} = (v_1, v_2)$ as follows:

$$\boldsymbol{v} = \int_\Theta (a \cos\beta(\theta), b \sin\beta(\theta))\, U(d\theta), \tag{4.1}$$

the random mean of Ferguson process on the ellipse. Before giving the probability density function of the random vector $\boldsymbol{v}$, we need to obtain its multivariate c-characteristic function. All proofs in this section are given by Jiang, Dickey, and Kuo (2003).

Lemma 4.1 *The c-characteristic function of $\boldsymbol{v}$, as defined in (4.1), has the closed form,*

$$g(\boldsymbol{t}; \boldsymbol{v}, c) = \left[\frac{2}{1 + \sqrt{1 + a^2 t_1^2 + b^2 t_2^2}}\right]^c, \quad \text{where } |\,\boldsymbol{t}\,| < (a^2 + b^2)^{-1/2}.$$

In the next lemma, we define a random vector of interest and find its multivariate c-characteristic function.

Lemma 4.2 *Let a random vector* $\boldsymbol{w}' = (w_1, w_2)$ *has the probability density function*

$$f(\boldsymbol{w}) = \frac{c}{\pi ab}\left(1 - \frac{w_1^2}{a^2} - \frac{w_2^2}{b^2}\right)^{c-1},$$

where a, b, *and* c *are positive real numbers and* $0 \le w_1^2/a^2 + w_2^2/b^2 < 1$. *Then the c-characteristic function of* $\boldsymbol{w}$ *is*

$$g(\boldsymbol{t}; \boldsymbol{w}, c) = \left[\frac{2}{1 + \sqrt{1 + a^2 t_1^2 + b^2 t_2^2}}\right]^c, \quad \text{where } |\,\boldsymbol{t}\,| < (a^2 + b^2)^{-1/2}.$$

Finally, by lemmas (3.2), (4.1), and (4.2), we are ready to give the distribution of $\boldsymbol{v}$, as a density in closed form, in the following theorem.

Theorem 4.3 *The probability density function of* $\boldsymbol{v}$*, which is defined in equation* (4.1)*, can be expressed as*

$$f(\boldsymbol{v}) = \frac{c}{\pi ab}\left(1 - \frac{v_1^2}{a^2} - \frac{v_2^2}{b^2}\right)^{c-1}, \tag{4.2}$$

where $0 \le v_1^2/a^2 + v_2^2/b^2 < 1$.

Theorem 4.3 says that the random functional of a Ferguson process, on an ellipse, has a probability density function (4.2). For the special case, when $a = b = 1$, we have that $\int_\Theta e^{i\beta(\theta)}\, U(d\theta)$ follows a distribution with probability density $f(\boldsymbol{v}) = (c/\pi)(1 - v_1^2 - v_2^2)^{c-1}$ on the unit disk, which is the same result given by Theorem 2 of Jiang (1991). The latter includes the elegant case of the uniform distribution on the disk for the random mean of the Ferguson process with parameter measure the uniform probability distribution on the circle.

5 Distribution of a random functional of a Ferguson process over the interval $[-1, 1]$

In this section, we shall give special one dimension random functionals of Ferguson processes. It can be seen that

$$\boldsymbol{X} = \int_\Theta e^{i\theta}\, U(d\theta) \tag{5.1}$$

is a spherical distribution. We shall study its marginal distribution. First, the following lemma can be shown by the definition of the random functional of a Ferguson's process.

Lemma 5.1 *Assume that* f *is a function that maps* y_1 *in* Y_1 *into* y_2 *in* Y_2, *i.e.,* $y_2 = f(y_1)$, *where* $y_1 \in Y_1$ *and* $y_2 \in Y_2$. *Let* V_1 *and* V_2 *be Ferguson processes with parameter measures* ν_1 *and* ν_2 *on* (Y_1, B_1) *and on* (Y_2, B_2), *respectively, where* B_i *is the* σ*-field of Borel subsets of* Y_i, *for* $i = 1, 2$. *We then have*

$$\int_{Y_1} f(y_1)\, V_1(dy_1) = \int_{Y_2} y_2\, V_2(dy_2),$$

if $\nu_2(s_2) = \nu_1(s_1)$ *and* $s_2 = \{\, y_2 \mid y_2 = f(y_1), y_1 \in s_1 \,\}$.

Next, we give the following marginal measure lemma.

Lemma 5.2 *Let μ be a uniform measure on the unit circle with center $(0,0)$ and total measure c, then its marginal measure μ_1 on X-axis is*

$$\mu_1[-1,x] = \frac{c}{\pi}\left(\arcsin x + \frac{\pi}{2}\right).$$

Proof. Let $x^2+y^2=1$ be the unit circle with center at $(0,0)$. For the semicircle above the X-axis, $y=\sqrt{1-x^2}$. Since $(dy)/(dx) = -x/\sqrt{1-x^2}$, it can be shown that the arc length of points A and B on this semicircle is

$$\int_a^b \sqrt{1+\left(\frac{dy}{dx}\right)^2} = \arcsin b - \arcsin a,$$

where a and b are the X coordinates of points A and B. Since μ is a uniform measure with total measure c, hence $\mu_1[-1,x] = (c/\pi)[\arcsin x - \arcsin(-1)] = (c/\pi)(\arcsin x + \pi/2)$. ❑

Finally, we give a one-dimensional random mean of the Ferguson process in the next theorem.

Theorem 5.3 *Let U_1 be a Dirichlet process with parameter μ_1 on $([-1,1],B_1)$, where B_1 is the σ-field of Borel subsets of $[-1,1]$ and $\mu_1[-1,x]=(c/\pi)(\arcsin x+\pi/2)$. Then the random mean of the Ferguson process*

$$X_1 = \int_{-1}^{1} x\, U(dx)$$

has the following probability density function

$$f(x_1) = \frac{1}{B(c+1/2,1/2)}(1-x_1^2)^{c-1/2}, \qquad -1<x_1<1.$$

Proof. By Corollary 3 of Jiang (1991),

$$(X_1,X_2) = \int_\Theta (\cos\theta, \sin\theta)\, U(d\theta)$$

has probability density function $f(x_1,x_2) = (c/\pi)(1-x_1^2-x_2^2)^{c-1}$, where $0 \le x_1^2+x_2^2 < 1$. Hence, $X_1 = \int_\Theta \cos\theta\, U(d\theta)$ has marginal probability density function $f_1(x_1) = (1/B(c+1/2,1/2))(1-x_1^2)^{c-1/2}$, $-1<x_1<1$. By Lemma 5.1 and Lemma 5.2,

$$\int_\Theta \cos\theta\, U(d\theta) = \int_{-1}^{1} x_1\, U_1(dx_1).$$

Therefore, $f_1(x_1)$ is also the probability density function of $\int_{-1}^{1} x_1\, U_1(dx_1)$. ❑

6 Conclusions

Multivariate c-characteristic function provides a new alternative method to deal with multivariate distributions that are hard to use the traditional characteristic function. It has many properties that are similar to the traditional characteristic function. With this new characteristic function, we expand the results that were given by Jiang (1988, 1991). In addition, we give two identical random functionals of two different Ferguson processes when certain conditions are met. With this result, we give the probability density function of a new one-dimensional random mean of a Ferguson process.

References

[1] Cifarelli, D. M., and Regazzini, E. (1990), "Distribution Functions of Means of a Dirichlet Process," Annals of Statistics, 18, 429-442.

[2] Diaconis, P., and Kemperman, J. (1996), "Some New Tools for Dirichlet Priors," in Bayesian Statistics 5, eds. J. M. Bernardo, J. O. Berger, A. P. Dawid, and A. F. M. Smith, New York: Oxford University Press, 97-106.

[3] Dickey, J. M. and Jiang, T. J. (1998), "Filtered-Variate Prior Distributions for Histogram Smoothing," Journal of the American Statistical Association, 93, 651–662.

[4] Ferguson, T. S. (1973), "A Bayesian analysis of some nonparametric problems," Ann. Statist., 1, 209–230.

[5] Hannum, R., Hollander, M., and Langberg, N. (1981), "Distributional results for random functionals of a Dirichlet process," Ann. Probab., 9, 665–670.

[6] Jiang, J. (1984), "Distribution Properties of Linear Forms in a Dirichlet Vector and Applications," Technical Report 7/84-14, State University of New York at Albany, Dept. of Mathematics and Statistics.

[7] Jiang, J. (1988), "Starlike functions and linear functions of a Dirichlet distributed vector," SIAM J. Math. Anal., 19, 390–397.

[8] Jiang, T. J. (1991), "Distribution of random functional of Dirichlet process on the unit disk," Statistics and Probability Letters, 12, 263–265.

[9] Jiang, T., Dickey, J. M., and Kuo, K. (2003), "A New Multivariate Transform and The Distribution of a Random Functional of a Ferguson Dirichlet Process." To appear in Stochastic Processes and Their Applications.

[10] Yamato, H. (1984), "Characteristic functions of means of distributions chosen from a Dirichlet process," Ann. Probab., 12, 262–267.

附件一　　　　出席國際學術會議心得報告　　　　姜志銘

2002的聯合統計會議，是由美國統計學會及其他好幾個國際知名的統計學會如加拿大統計學社、IMS、ENAR、WNAR等，於2000年8月11–15日，在前一年911事件發生地的美國紐約市舉行，雖然地點敏感（離911事件主要發生地世貿大樓不遠）且911事件即將屆滿一週年，但來自全球各地的統計專家、學者等仍有幾千人為追求最新的學術新知而參加。

本次會議的主題為statistics in an era of technological change，各種領域的會議都排得很緊湊，收穫相當多，但其中印象最深刻的是在8月13日（星期二）黃昏的Deming Lecture，Deming教授為一已去世的統計大師，尤其第二次世界大戰後，教育日本工業界的統計品管，使得日本產品的品質能很快得到肯定，一般相信，日本能於戰後很快的變成經濟強國，Deming教授的功勞相當大。在這一個節目中，在介紹完Deming教授對統計社會及整個社會的貢獻後，接下來的是英國爵士D. R. Cox教授先接受由美國統計學會會長頒發的獎牌，以及他的演講，此時台下同時至少有千人以上的現象，聆聽這位在統計各領域都有傑出成就的統計大師專題演講的風采。我的演講題目為On a new characteristic function（本文主要是探討並提供新的特徵函數，它的一些特性、應用，以及它與傳統特徵函數的關係等），排在星期四Recent Advances in Estimation session。雖然排到最後一天，但來聽演講的仍不少，有些觀眾甚感興趣，甚至在會後來向我要了論文。

參加這次會議的另一項收穫是遇見我的博士論文指導教授Jim Dickey，因為我們仍一直有論文方面的合作，因此我們也趁著這個機會，見了很多次面，花了很長的時間討論，一直合作以及未來可能合作的論文。

最後謝謝國科會給予參加這次非常忙碌，但收穫相當豐富的會議。

出席國際學術會議發表之論文

On a new characteristic function

Thomas J. Jiang and Kun-Lin Kuo
Department of Mathematical Sciences
National Chengchi University
#64 Chih-Nan Road, Section 2
Taipei 11623
TAIWAN

Abstract

A new multivariate c-characteristic function has been shown to have uniqueness and convergence properties, which are similar to those of the regular characteristic function. In this paper, we shall give further properties of the multivariate c-characteristic function. We shall also give an example to show how the probability density function can be found when its corresponding new characteristic function is known. The relation between this new characteristic function and the traditional characteristic function shall also be discussed.

Keywords: characteristic function, Dirichlet distribution, Carlson's function, inversion formula, filtered-variate Dirichlet distribution.