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

過濾變量分配及其貝氏計算

On the filtered-variate distribution and their Bayesian computations

計畫類別：■個別型計畫 □整合型計畫
計畫編號：NSC 94-2118-M-004-006
執行期間：94年8月1日至95年7月31日

計畫主持人：姜志銘
計畫參與人員：郭錕霖 林其緯 李婉菁 林書淵

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 五 年 十 月 二 十 四 日

行政院國家科學委員會專題研究計畫成果報告

過濾變量分配及其貝氏計算

On the filtered-variate distribution and their Bayesian computations

計畫編號：NSC 94-2118-M-004-006

執行期限：94年8月1日至95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　林其緯　李婉菁　林書淵

國立政治大學應用數學系

中文摘要

Dirichlet分配群為多項分配抽樣的自然共軛家族。然而，鄰近類別的抽樣機率值通常相當接近，因此具有平滑性的先驗分配家族就有很高的需求性。不幸的是，Dirichlet分配的鄰近類別具有負相關的特性，使其本身不具平滑性，為了解決這樣的問題，Dickey and Jiang (1998)首先提出過濾變量分配，特別是當「標準的」Dirichlet隨機機率向量經過適當的線性轉換，亦即過濾變量的Dirichlet分配，可被證明具有平滑性。

此處，我們將先給一個對研究過濾變量的Dirichlet分配相當重要的一般性c-特徵函數的反轉公式。接下來，我們將提供單維的過濾變量Dirichlet分配的機率密度函數，這對後續先驗平滑性的研究將有相當的幫助。更多的結果可參考Jiang and Kuo (2006a,b)。

關鍵詞：過濾變量的Dirichlet分配，平滑性，c-特徵函數。

Abstract

The class of Dirichlet distributions is a natural conjugate prior family for the multinomial sampling. However, the sampling probabilities with adjacent categories are likely to be close in value. It is therefore highly desirable to have "smooth" prior distributions family. Unfortunately, the Dirichlet distribution have the property where corresponding adjacent random quantities have negative correlations, which are in violation of the smoothness assumption. To deal with such a problem, Dickey and Jiang (1998) define the filtered-variate distribution. In particular, the filtered-variate Dirichlet distribution, with a suitable linear transform of a "standard" Dirichlet random probability vector, can be shown to be smooth.

Here, we first give a general inversion formula of the c-characteristic function, which is an important tool to study the filtered-variate Dirichlet distribution. We then give the probability density function of the one-dimensional distribution of the filtered-variate Dirichlet distribution, which is potentially important in helping constructing the smooth prior in future research. Additional results are given in Jiang and Kuo (2006a,b).

Keywords: filtered-variate Dirichlet distribution, smoothness, c-characteristic function

1 Introduction

The class of Dirichlet distributions is a natural conjugate prior family for the multinomial sampling. However, the sampling probabilities with adjacent categories are likely to be close in value. It is therefore highly desirable to have "smooth" prior distributions family. Unfortunately, the Dirichlet distribution have the property where corresponding adjacent random quantities have negative correlations, which are in violation of the smoothness assumption. To deal with such a problem, Dickey and Jiang (1998) define the filtered-variate distribution. In particular, the filtered-variate Dirichlet distribution, with a suitable linear transform of a "standard" Dirichlet random probability vector, can be shown to be smooth.

The filtered-variate Dirichlet distribution is easily characterized by the c-characteristic function, which can be seen in Jiang, Dickey, and Kuo (2004). The usefulness of the c-characteristic function would be hindered without its inversion formula. Jiang and Kuo (2006a) first give a version of the c-characteristic function with the restriction of $c = 1$. In the next section, we shall first give a general version of the inversion formula of the univariate c-characteristic function. With this general inversion formula, there are many applications, e.g., random functionals of Ferguson-Dirichlet processes, can be found. Here, we shall use a special case of this inversion formula to find the linear combination of the Dirichlet random vector, which is the one dimensional case of the filtered-variate Dirichlet distribution. Conclusions are given in the last section.

2 Inversion formula of the c-characteristic function and its application to the filtered-variate Dirichlet distribution

Before giving an inversion formula of the c-characteristic function, we state the c-characteristic function's definition and some of its properties first.

Definition 1 *For any finite measure μ with supports in $[-a, a]$, we define the c-transformation of μ as*

$$\hat{\mu}^c(t) = \int_{-a}^{a} (1 - itx)^{-c}\, d\mu(x), \ \ |t| < a^{-1}, \ \ c > 0,$$

where a is a positive real number.

The following lemma shows that there is a one-to-one correspondence between $\hat{\mu}^c(t)$ and μ.

Lemma 2 (Jiang, 1988) *For any finite measures μ and ν with supports in $[-a, a]$ and any positive real number c, if we have*

$$\hat{\mu}^c(t) = \hat{\nu}^c(t),$$

for all $|t| < a^{-1}$, then $\mu = \nu$.

Jiang (1988) shows the following convergence theorem.

Theorem 3 (Jiang, 1988) *Given c, assume that sub-probability measures* $\mu, \mu_1, \mu_2, \ldots$ *(with supports in* $[-a, a]$*) correspond to* $\hat{\mu}^c(t), \hat{\mu}_1^c(t), \hat{\mu}_2^c(t), \ldots$*, respectively. If for all* $|t| < a^{-1}$

$$\hat{\mu}_n^c(t) \to \hat{\mu}^c(t) \text{ as } n \to \infty,$$

then

$$\mu_n \xrightarrow{v} \mu \text{ as } n \to \infty.$$

For any random variable X, we define its c-characteristic function, denoted by $g(t, X, c)$, as $\hat{\mu}^c(t)$ if its corresponding probability measure is μ. That is, $g(t; X, c) = \hat{\mu}^c(t)$. Given a sequence of random variables $X_1, X_2, \ldots$, with its corresponding sequence of c-characteristic functions $g(t; X_1, c)$, $g(t; X_2, c)$, $\ldots$, it is very possible that we may be able to find the limiting distribution of $g(t; X_n, c)$. If there is a random variable X having corresponding c-characteristic function $g(t; X, c)$, and $\lim_{n\to\infty} g(t; X_n, c) = g(t; X, c)$, for all $|t| < a^{-1}$. By Lemma 2 and Theorem 3, then X_n converges in distribution to X. In practice, it may not be easy to guess what the random variable X having c-characteristic function $g(t; X, c)$ is. In the next section, we shall provided methods to find the probability density function of X when its c-characteristic function is available.

First, we give a version of the inversion formula of the generalized Stieltjes transformation given by Sumner (1949).

Theorem 4 (Sumner, 1949, Theorem 4a) *If c is positive and* $f \in L[0, b]$ *for all positive b, and is such that* $h(t) = \int_0^\infty (x+t)^{-c} f(x)\, dx$ *converges, then*

$$\lim_{\epsilon\to 0^+} \frac{-1}{2\pi i} \int_{C_{\epsilon,x}} (z+x)^{c-1} h'(z)\, dz = \frac{f(x^+) + f(x^-)}{2}$$

for any positive x at which both $f(x^+)$ *and* $f(x^-)$ *exist, where* $C_{\epsilon,x}$ *is the contour which starts at the point* $-x - i\epsilon$*, proceeds along the straight line* $Im\, z = -\epsilon$ *to the point* $-i\epsilon$*, then along the semi-circle* $|z| = \epsilon$*,* $Re\, z \geq 0$*, to the point* $i\epsilon$*, and finally along the line* $Im\, z = \epsilon$ *to the point* $-x + i\epsilon$*.*

We give another version of inverse formula for the generalized Stieltjes transform in the following corollary.

Corollary 5 *If* $a \in R$*, c is positive and* $f(x) \in L[a, b]$ *for all* $b > a$*, and is such that the integral*

$$h(t) = \int_a^\infty (x+t)^{-c} f(x)\, dx$$

converges, then

$$\frac{f(x^+) + f(x^-)}{2} = \lim_{\epsilon\to 0^+} \frac{-1}{2\pi i} \int_{D_{\epsilon,x}} (z+x)^{c-1} h'(z)\, dz$$

for any $x > a$ *at which both* $f(x^+)$ *and* $f(x^-)$ *exist, where* $D_{\epsilon,x}$ *is the contour which starts at the point* $-x - i\epsilon$*, proceeds along the straight line* $Im\, z = -\epsilon$ *to the point* $-a - i\epsilon$*, then along the semi-circle* $|z + a| = \epsilon$*,* $Re\, z \geq -a$*, to the point* $-a + i\epsilon$*, and finally along the line* $Im\, z = \epsilon$ *to the point* $-x + i\epsilon$*.*

We then have the following inversion formula for our c-characteristic function.

Theorem 6 *If $g(t;X,c) = \int_a^b (1-itx)^{-c} f(x)\,dx$ converges, then*

$$\begin{aligned}
&\frac{f(x^+)+f(x^-)}{2} \\
&= \lim_{\epsilon\to 0^+} \frac{1}{2\pi i}\left\{\int_a^x \left[(x-y+i\epsilon)^{c-1}\left.\frac{dG(t)}{dt}\right|_{t=-y+i\epsilon} - (x-y-i\epsilon)^{c-1}\left.\frac{dG(t)}{dt}\right|_{t=-y-i\epsilon}\right] dy \right. \\
&\quad \left. - \int_{-\pi/2}^{\pi/2} (-a+x+\epsilon e^{i\theta})^{c-1}\left.\frac{dG(t)}{dt}\right|_{t=-a+\epsilon e^{i\theta}} i\epsilon e^{i\theta}\,d\theta\right\},
\end{aligned}$$

where $G(t) = t^{-c} g(i/t;X,c)$, for any $x \in (a,b)$ at which both $f(x^+)$ and $f(x^-)$ exist.

The following corollary is a special case of Theorem 6, when $c = 1$.

Corollary 7 *If $g(t;X,1) = \int_a^b (1-itx)^{-1} f(x)\,dx$ converges, then*

$$\frac{f(x^+)+f(x^-)}{2} = \lim_{\epsilon\to 0^+}\frac{1}{2\pi i}\left[\frac{g\left(\frac{i}{-x-i\epsilon};X,1\right)}{-x-i\epsilon} - \frac{g\left(\frac{i}{-x+i\epsilon};X,1\right)}{-x+i\epsilon}\right],$$

for any $x \in (a,b)$ at which both $f(x^+)$ and $f(x^-)$ exist.

The above Corollary 7 can also be obtained directly, see Jiang and Kuo (2006a), by Stieltjes transform.

Let $\boldsymbol{u} = (u_1,\ldots,u_I)$ be a Dirichlet distribution with parameters $\boldsymbol{b} = (b_1,\ldots,b_I)$, denoted by $\boldsymbol{u} \sim D(\boldsymbol{b})$, then the probability density function of $\boldsymbol{u}$ can be expressed as $[B(\boldsymbol{b})]^{-1}\prod_{i=1}^I u_i^{b_i-1}$, where $u_i > 0$, $u_+ = \sum_{i=1}^I u_i = 1$, and $B(\boldsymbol{b}) = \prod_{i=1}^I \Gamma(b_i)/\Gamma(b_+)$. In the next example, we shall apply the above inversion formula to find the probability density function of the linear combination of the Dirichlet random variables, say $Y = \sum_{i=1}^I w_i u_i$, when $c = b_+ = 1$. Without loss of the generality, we shall consider the case that $w_1 < w_2 < \cdots < w_I$ only.

Example 1 Let $\boldsymbol{u} = (u_1,\ldots,u_I) \sim D(\boldsymbol{b})$, where $b_+ = 1$. Consider a new random variable $Y = \sum_{i=1}^I w_i u_i$, where $w_1 < w_2 < \cdots < w_I$. Then, for $w_m < y < w_{m+1}$ and $m = 1,2,\ldots$, or $I-1$, the probability density function $g(y)$ of Y can be expressed as

$$g(y) = \frac{1}{\Gamma(\sum_{j=1}^m b_j)\Gamma(\sum_{j=m+1}^I b_j)}\prod_{j=1}^m (y-w_j)^{-b_j}\prod_{j=m+1}^I (w_j-y)^{-b_j}.$$

Solution. The c-characteristic function of Y, by Corollary 3.3 of Jiang (1988), is $g(t;Y,1) = \prod_{j=1}^I (1-itw_j)^{-b_j}$. We shall use Corollary 7 to find the probability density function of Y. First, we have

$$\begin{aligned}
&\lim_{\epsilon\to 0^+}\frac{1}{2\pi i}\left[\frac{g\left(\frac{i}{-y-i\epsilon};Y,1\right)}{-y-i\epsilon} - \frac{g\left(\frac{i}{-y+i\epsilon};Y,1\right)}{-y+i\epsilon}\right] \\
&= \lim_{\epsilon\to 0^+}\frac{1}{2\pi i}\left[\prod_{j=1}^I (-y+w_j-i\epsilon)^{-b_j} - \prod_{j=1}^I (-y+w_j+i\epsilon)^{-b_j}\right].
\end{aligned}$$

We also have

$$\begin{aligned}(-y+w_j-i\epsilon)^{-b_j} &= \exp[-b_j\ln(-y+w_j-i\epsilon)] \\ &= \exp[-b_j\ln|-y+w_j-i\epsilon|]\exp[-ib_j\arg(-y+w_j-i\epsilon)] \\ &= |-y+w_j-i\epsilon|^{-b_j}\exp[-ib_j\arg(-y+w_j-i\epsilon)].\end{aligned}$$

Given $m=1,2,\ldots,$ or $I-1$, when $w_m<y<w_{m+1}$, then $-y+w_j<0$ for all $j\le m$ and $-y+w_j>0$ for all $j\ge m+1$. Hence, when $w_m<y<w_{m+1}$, we have

$$\lim_{\epsilon\to0^+}(-y+w_j-i\epsilon)^{-b_j}=\begin{cases}(y-w_j)^{-b_j}\exp[ib_j\pi], & \text{when } j=1,2,\ldots,m,\\ (w_j-y)^{-b_j}, & \text{when } j=m+1,m+2,\ldots,I.\end{cases}$$

Similarly, when $w_m<y<w_{m+1}$, we have

$$\lim_{\epsilon\to0^+}(-y+w_j+i\epsilon)^{-b_j}=\begin{cases}(y-w_j)^{-b_j}\exp[-ib_j\pi], & \text{when } j=1,2,\ldots,m,\\ (w_j-y)^{-b_j}, & \text{when } j=m+1,m+2,\ldots,I.\end{cases}$$

The probability density function of Y, $g(y)$, when $w_m<y<w_{m+1}$, can be expressed as

$$\begin{aligned}g(y) &= \frac{1}{2\pi i}\left\{\prod_{j=1}^{m}(y-w_j)^{-b_j}\prod_{j=m+1}^{I}(w_j-y)^{-b_j}\exp\left[i\left(\sum_{j=1}^{m}b_j\right)\pi\right]\right. \\ &\quad\left.-\prod_{j=1}^{m}(y-w_j)^{-b_j}\prod_{j=m+1}^{I}(w_j-y)^{-b_j}\exp\left[-i\left(\sum_{j=1}^{m}b_j\right)\pi\right]\right\} \\ &= \frac{1}{2\pi i}\left\{\prod_{j=1}^{m}(y-w_j)^{-b_j}\prod_{j=m+1}^{I}(w_j-y)^{-b_j}2i\sin\left[\left(\sum_{j=1}^{m}b_j\right)\pi\right]\right\} \\ &= \frac{1}{\Gamma(\sum_{j=1}^{m}b_j)\Gamma(\sum_{j=m+1}^{I}b_j)}\prod_{j=1}^{m}(y-w_j)^{-b_j}\prod_{j=m+1}^{I}(w_j-y)^{-b_j}.\end{aligned}$$

The last identity can be obtained by using Theorem 3.9-1 of Carlson (1977, p. 48). ■

3 Conclusions

The c-characteristic function and its inversion formula are important tools to study the filtered-variate Dirichlet distribution. Jiang and Kuo (2006b) give additional proofs and example on applying the inversion formula. Here, we give an easy way to find the probability density function of the linear combination of the Dirichlet random probability vector. This is potentially important in helping constructing the smooth prior in future research.

References

Carlson, B. C. (1977), *Special Functions of Applied Mathematics*, Academic Press, New York.

Dickey, J. M. and Jiang, T. J. (1998), "Filtered-variate prior distributions for historgram smoothing," *Journal of the American Statistical Association*, **93**, 651–662.

Jiang, J. (1988), "Starlike functions and linear functions of a Dirichlet distributed vector," *SIAM Jouurnal on Mathematical Analysis*, **19**, 390–397.

Jiang, T.J., Dickey, J.M., and Kuo, K.-L. (2004), "A new multivariate transform and the distribution of a random functional of a Ferguson-Dirichlet process," *Stochastic Processes and their Applications*, **111**, 77–95.

Jiang, T.J. and Kuo, K.-L. (2006a), "An inverse formula of the univariate c-chracteristic function," *Proceedings of 5th Annual Hawaii International Conference on Statistics, Mathematics and Related Fields*, 994–998.

Jiang, T.J. and Kuo, K.-L. (2006b), "On the ranom functional of the Ferguson-Dirichlet process." To be appeared in *Proceeding of the Section on Bayesian Statistical Science of the American Statistical Association.*

Sumner, D.B. (1949), "An inversion formula for the generalized Stieltjes transform," *Bulletin of the American Mathematical Society*, **55**, 174–183.