

非対称な円周分布による有限混合分布と その推定について

高崎経済大学・経済学部* 宮田 庸一

Yoichi Miyata

Faculty of Economics,

Takasaki City University of Economics

東京理科大学・工学部† 塩濱 敬之

Takayuki Shiohama

Department of Information and Computer Technology,

Tokyo University of Science

南山大学・理工学部‡ 阿部 俊弘

Toshihiro Abe

Faculty of Science and Engineering, Nanzan University

概要

本稿では、円周上の非対称な確率分布、およびそれらの有限混合モデルについて紹介する。特に非対称な確率分布の有限混合モデルにおける最尤推定量を求めるためのアルゴリズムおよび推定量の一致性に関する結果を紹介する。

1 はじめに

0度と360度が同じことを意味する角度のデータにおいては、実軸上の数値と同様の統計解析を行うと不具合が生じる可能性があることが知られている。このようなデータは、

* 〒370-0801 高崎市上並榎町 1300

† 〒125-8585 東京都葛飾区新宿 6-3-1

‡ 〒466-8673 名古屋市昭和区山里町 18

古典地磁気学, 生物学, 環境学等様々な分野で存在し, それを説明するために, 単位円周上, もしくは, より一般的に多様体上の確率モデルを考える必要がある. 単位円周上の確率モデルとして, 有名なものは von Mises 分布と wrapped Cauchy 分布である. これらはいずれも対称な分布であるが, 実際のデータは非対称であることや, 複数のモードを持つこともしばしばある. 非対称な分布を得る手法としては, 基準となる対称分布の正弦摂動によって非対称性を表現する手法 (Azzalini and Capitanio (2003) や Abe and Pewsey (2011)) と, 基準となる対称分布の尺度変換を利用した手法 (Jones and Pewsey (2012)) がある. 一方で非対称性もしくは多峰性に対応するため Fraser et al. (1981), Holzmann et al. (2004), Banerjee et al. (2005) 等は, 対称な円周分布を用いた有限混合モデルについて研究を行った. 本稿においては, 円周上の基本的な確率モデルを紹介した後に, 非対称な分布への拡張, および非対称な分布を用いた有限混合モデルについて紹介を行う. また非対称な分布の有限混合モデルにおいて最尤推定を行うための EM アルゴリズムおよびその強一緻性に関する結果についても紹介する.

2 基本統計量と分布の定義

2.1 円周上の確率密度関数

f が円周分布の確率密度関数 (pdf) であるとは次の 3 つが成立することをいう.

- (a) $f(\theta) \geq 0 \quad a.e. \theta \in (-\infty, \infty)$,
- (b) $f(\theta + 2\pi) = f(\theta) \quad a.e. \theta \in (-\infty, \infty)$,
- (c) $\int_{-\pi}^{\pi} f(\theta) d\theta = 1$.

角度の確率変数 Θ の p 次三角モーメント (trigonometric moment) は

$$\phi_p = E(e^{ip\Theta}), \quad i^2 = -1, \quad p = 1, 2, 3, \dots$$

で定義され, $\phi_1 = E(e^{i\Theta}) = \rho e^{i\mu}$ のときの μ を平均方向 (mean direction), ρ を平均合成ベクトル長 (mean resultant length) という. また, ϕ_p の実部と虚部はそれぞれ, p 次の cosine モーメントと p 次の sine モーメントといい,

$$\alpha_p = E(\cos(p\Theta)), \quad \beta_p = E(\sin(p\Theta))$$

により与えられる.

2.2 平均方向と集中度

角度データ $\theta_1, \dots, \theta_n$ に対して,

$$\bar{C} = \frac{1}{n} \sum_{j=1}^n \cos \theta_j, \quad \bar{S} = \frac{1}{n} \sum_{j=1}^n \sin \theta_j$$

とおくと、標本平均合成ベクトル長 (sample mean resultant length) は

$$\bar{R} = \sqrt{\bar{C}^2 + \bar{S}^2}$$

で与えられる。標本平均方向 (sample mean direction) は $\bar{R} > 0$ ならば,

$$\bar{\theta} = \arg(\bar{C} + i\bar{S})$$

として定義される。 $\bar{R} = 0$ のとき、 $\bar{\theta}$ は定義されない。

標本平均合成ベクトル長 \bar{R} は $0 \leq \bar{R} \leq 1$ を満たす。もし、 $\theta_1, \dots, \theta_n$ が同じような方向に集まっているならば、 \bar{R} は 1 に近づき、 $\theta_1, \dots, \theta_n$ が広く散らばっているならば、 \bar{R} は 0 に近づく。この意味で、 \bar{R} はデータの「集中度 (concentration)」を測っていることがわかる。注意点として、 \bar{R} が 0 に近くても、 $\theta_1, \dots, \theta_n$ が広く散らばっていない場合がある：集中度がある一方で、散らばり具合を測るものとして、「円周分散 (circular variance)」は

$$V = 1 - \bar{R}$$

により定義される。これも $0 \leq V \leq 1$ を満たす。

\bar{C} と \bar{S} を拡張したものとして、0 周りの p 次の標本三角モーメントが定義される:

$$a_p = \frac{1}{n} \sum_{j=1}^n \cos p\theta_j, \quad b_p = \frac{1}{n} \sum_{j=1}^n \sin p\theta_j.$$

平均方向 $\bar{\theta}$ 周りでは

$$\bar{a}_p = \frac{1}{n} \sum_{j=1}^n \cos p(\theta_j - \bar{\theta}), \quad \bar{b}_p = \frac{1}{n} \sum_{j=1}^n \sin p(\theta_j - \bar{\theta})$$

により定義される。

3 円周上の対称分布

3.1 一様分布

最も基本的な円周分布は (円周) 一様分布である. この分布はしばしば, 角度データの出現頻度の一様性を検定する仮説検定の帰無分布としても用いられ, その確率密度関数は

$$f(\theta) = \frac{1}{2\pi}, \quad -\pi \leq \theta < \pi$$

である. 一様分布の p 次三角モーメントは

$$\phi_p = \begin{cases} 1, & p = 0, \\ 0, & p \neq 0 \end{cases}$$

により与えられる.

3.2 Von Mises 分布

円周分布でおそらく一番良く知られている分布は von Mises 分布 (von Mises (1918)) である. この分布は (a) 二変量正規分布を極座標変換し, 長さ一定の下での条件付分布として得られること, (b) 位置パラメータの最尤推定量が標本平均方向であるという特徴付けから導けることや (c) 一次の cosine モーメントと一次の sine モーメントが一定という仮定の下でエントロピーを最大化することにより得られること等から, 円周上の正規分布 (circular normal distribution) と呼ばれている (Jammalamadaka and SenGupta (2001)). しかし, von Mises 分布は和に関する再生性を持たないこと等もあり, この呼び方は研究者によって意見が分かれている.

von Mises 分布 $VM(\mu, \kappa)$ の確率密度関数は

$$f(\theta) = \frac{1}{2\pi I_0(\kappa)} \exp[\kappa \cos(\theta - \mu)], \quad -\pi \leq \theta < \pi$$

で与えられる. ここで, $\mu \in [-\pi, \pi)$ は位置パラメータ, $\kappa \in [0, \infty)$ は集中パラメータである. I_p は p 次の第 1 種変形 Bessel 関数であり,

$$I_p(\kappa) = \frac{1}{2\pi} \int_0^{2\pi} \cos p\theta e^{\kappa \cos \theta} d\theta = \sum_{r=0}^{\infty} \frac{1}{\Gamma(p+r+1)r!} \left(\frac{\kappa}{2}\right)^{2r+p}$$

で与えられる. 集中パラメータ κ の範囲は負にしても確率密度関数になっているが, $VM(\mu, -\kappa) = VM(\mu + \pi, \kappa)$ であることとパラメータの解釈から通常は $\kappa \geq 0$ で定義さ

れる. このようなことから von Mises 分布に限らず, 他の多くの分布でも集中パラメータは通常, 非負である.

von Mises 分布の p 次三角モーメントが

$$\phi_p = \frac{I_p(\kappa)}{I_0(\kappa)} e^{ip\mu}, \quad p = \pm 1, \pm 2, \dots$$

であることから, 平均合成ベクトル長は $A(\kappa) = \frac{I_1(\kappa)}{I_0(\kappa)}$ である. また, $\theta_1, \dots, \theta_n$ に対して, 尤度関数 $\ell(\mu, \kappa)$ は,

$$\ell(\mu, \kappa) = \kappa \sum_{i=1}^n \cos(\theta_i - \mu) - n \log(2\pi I_0(\kappa))$$

で与えられ, 最尤推定量は $\hat{\mu} = \arg(\bar{C} + i\bar{S})$, $\hat{\kappa} = A^{-1}(\bar{R})$ となる. 一方で, von Mises 分布以外で最尤推定量を単純な形で与えられる分布は知られていない.

3.3 Wrapped Cauchy 分布

von Mises 分布の他に, よく知られている分布としては, wrapped Cauchy 分布 $WC(\mu, \rho)$ がある. wrapped Cauchy 分布は (実軸上の) Cauchy 分布に巻き込み法 (wrapping) を適用することにより得られることから, このような名前がつけられている. wrapped Cauchy 分布の確率密度関数は

$$f(\theta) = \frac{1}{2\pi} \frac{1 - \rho^2}{1 + \rho^2 - 2\rho \cos(\theta - \mu)}, \quad -\pi \leq \theta < \pi$$

により与えられる. ここで, $\mu \in [-\pi, \pi)$ は位置パラメータ, $\rho \in [0, 1)$ は集中パラメータである. wrapped Cauchy 分布の p 次三角モーメントは

$$\phi_p = \rho^{|p|} e^{ip\mu}, \quad p = \pm 1, \pm 2, \dots$$

であり, 平均合成ベクトル長は ρ である.

von Mises 分布は再生性を持たないが, wrapped Cauchy 分布は $WC(\mu_1, \rho_1)$ と $WC(\mu_2, \rho_2)$ に従う確率変数の和の分布が $WC(\mu_1 + \mu_2, \rho_1 \rho_2)$ となり, 再生性を持っている.

3.4 Cardioid 分布

一様分布を cosine 関数を用いて摂動すると cardioid 分布 $C(\mu, \rho)$ が得られ、その確率密度関数は

$$f(\theta) = \frac{1}{2\pi} (1 + 2\rho \cos(\theta - \mu)), \quad -\pi \leq \theta < \pi$$

により与えられる (例えば, Jeffreys (1948) を見よ). ここで, $\mu \in [-\pi, \pi)$ は位置パラメータ, $\rho \in [0, 1/2)$ は集中パラメータである. 極座標において, $r = f(\theta)$ とすると cardioid 曲線になることがこの分布の名前の由来である.

p 次三角モーメントは

$$\phi_p = \begin{cases} 1, & p = 0, \\ \rho e^{i\mu}, & p = \pm 1, \\ 0, & p = \pm 2, \pm 3, \dots \end{cases}$$

となる. この分布も $C(\mu_1, \rho_1)$ と $C(\mu_2, \rho_2)$ に従う確率変数の和の分布が $C(\mu_1 + \mu_2, \rho_1 \rho_2)$ となり, 再生性を持っている.

3.5 Jones–Pewsey 分布

Jones and Pewsey (2005) は既存の対称分布を含むような柔軟な対称分布族を提案した. その確率密度関数は

$$f(\theta) = \frac{(\cosh(\kappa\psi) + \sinh(\kappa\psi) \cos(\theta - \mu))^{1/\psi}}{2\pi P_{1/\psi}(\cosh(\kappa\psi))}, \quad -\pi \leq \theta < \pi \quad (1)$$

である. ここで, $\mu \in [-\pi, \pi)$ は位置パラメータ, $\kappa \in [0, \infty)$ は集中パラメータ, $\psi \in \mathbb{R}$ は形状パラメータ, $P_{1/\psi}$ は Legendre の陪関数であり,

$$P_\nu(z) = \frac{1}{\pi} \int_0^\pi (z + \sqrt{z^2 - 1} \cos x)^\nu dx$$

を満たす. この分布は 3 つのパラメータだけで前に述べた分布や他の対称分布を特別な場合として含んでいる: (1) で, $\psi \rightarrow 0$ としたとき, von Mises 分布, $\psi = 1$ のとき, cardioid 分布, $\psi = -1$ のとき, wrapped Cauchy 分布, $\psi > 0$ で $\kappa \rightarrow \infty$ としたとき, Cartwright's power-of-cosine 分布 (Cartwright (1963)) に帰着する. Jones–Pewsey 分布は Shimizu and Iida (2002) の円周 t 分布においてパラメータの範囲を広げたものと解釈することもできる (清水 (2008)). Jones–Pewsey 分布に関連した対称分布として, Abe et al. (2010) は投影図法を用いることにより, Jones–Pewsey 分布を含む対称分布族を提案している.

4 円周上の非対称分布について

実際のデータ解析において、非対称分布を適用することの利点が多いことから、種々の円周上の非対称分布も提案されている。本稿では、非対称性を取り入れる方法を2つ紹介する。一つ目として、基準となる対称分布の密度関数を $f_0(\theta)$ とすると、次のような確率分布を考えることで非対称性を表現することができる。

$$f(\theta - \mu) = (1 + \lambda \sin(\theta - \mu))f_0(\theta - \mu)$$

パラメータ $\lambda \in [-1, 1]$ が非対称性を表す指標であり、 $\lambda > 0$ (< 0) なら右 (左) に歪んだ分布を表す。 $\lambda = 0$ の場合は、基準となる対称分布となる。位置パラメータ μ は $\theta \mapsto \theta - \mu$ として導入される。この分布の p 次の三角モーメントは

$$\phi_p = \alpha_{0,p} + i \frac{\lambda(\alpha_{0,p-1} - \alpha_{0,p+1})}{2}, \quad p = \pm 1, \pm 2, \dots$$

となる。ここで、 $\alpha_{0,p}$ は確率密度関数が f_0 である参照する対称分布の p 次の cosine モーメントである。

基準となる分布を von Mises 分布とした場合、von-Mises 分布の正弦摂動による確率分布を Sine-skewed von Mises (SSvM) 分布といい、その密度関数は次のようになる。

$$f_{SSvM}(\theta|\mu, \kappa, \lambda) = \frac{1}{2\pi I_0(\kappa)} \exp\{\kappa \cos(\theta - \mu)\} \{1 + \lambda \sin(\theta - \mu)\},$$

ここで、 $0 \leq \mu < 2\pi$ は平均方向のパラメータで、 $\kappa \geq 0$ は集中度を表すパラメータである。式中の関数 $I_\nu(\kappa) := \frac{1}{2\pi} \int_0^{2\pi} \cos(\nu\theta) \exp\{\kappa \cos \theta\} d\theta$ は、次数 $\nu \in \mathbb{Z}$ の第一種修正ベッセル関数である。図1には、von Mises 分布の正弦摂動の密度関数を異なる λ に対してプロットしたものである。図より、集中度が高いと非対称性の影響は顕著に確認することはできないが、 $\kappa = 1$ の場合は、 λ の値を -1 に近づけることで左に歪んだ分布が表現できることがわかる。

von Mises 分布の正弦摂動の三角モーメントは次のように陽に与えられる。 $\alpha_p := I_p(\kappa)/I_0(\kappa)$ とおくと、

$$\begin{aligned} \alpha_p^* &:= E\{\cos(p\Theta)\} = \left\{ \cos(p\mu) - \frac{p\lambda}{\kappa} \sin(p\mu) \right\} \alpha_p, \\ \beta_p^* &:= E\{\sin(p\Theta)\} = \left\{ \sin(p\mu) + \frac{p\lambda}{\kappa} \cos(p\mu) \right\} \alpha_p. \end{aligned}$$

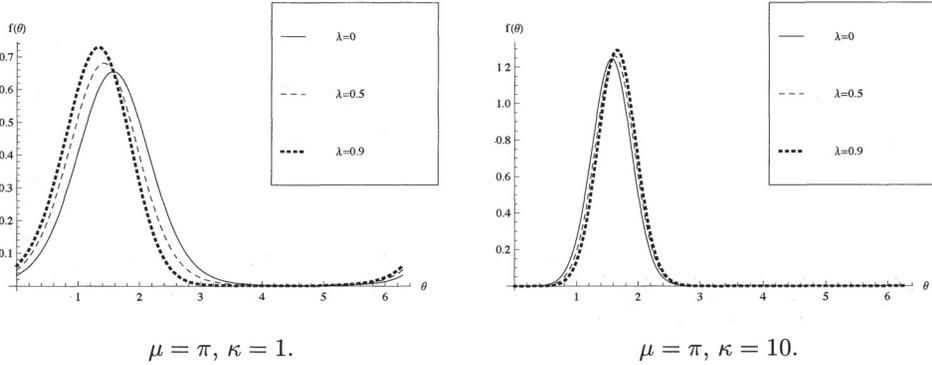


図1 von Mises 分布の正弦摂動の密度関数のプロット

これより、平均合成ベクトル長は $\rho = I_1(\kappa)\sqrt{\kappa^2 + \lambda^2}/(\kappa I_0(\kappa))$ となる。また、Abe and Pewsey (2011) より、SSvM 分布の単峰性は保証されない。

基準となる対称分布を巻き込みコーシー分布にすると、巻き込みコーシー分布の正弦摂動の密度関数が得られる (Sine-skewed Wrapped Cauchy (SSWC) 分布)。確率変数 Θ がパラメータ μ, ρ, λ の SSWC 分布に従うとき、 Θ の密度関数は次のようになる。

$$f_{SSWC}(\theta|\mu, \rho, \lambda) = \frac{1 - \rho^2}{2\pi(1 + \rho^2 - 2\rho\cos(\theta - \mu))} \{1 + \lambda\sin(\theta - \mu)\}, \quad (2)$$

ここで $\lambda \in [-1, 1]$ は非対称性のパラメータで $\lambda = 0$ のときに、巻き込みコーシー分布になる。図2には、巻き込みコーシー分布の正弦摂動の密度関数をプロットした。巻き込み

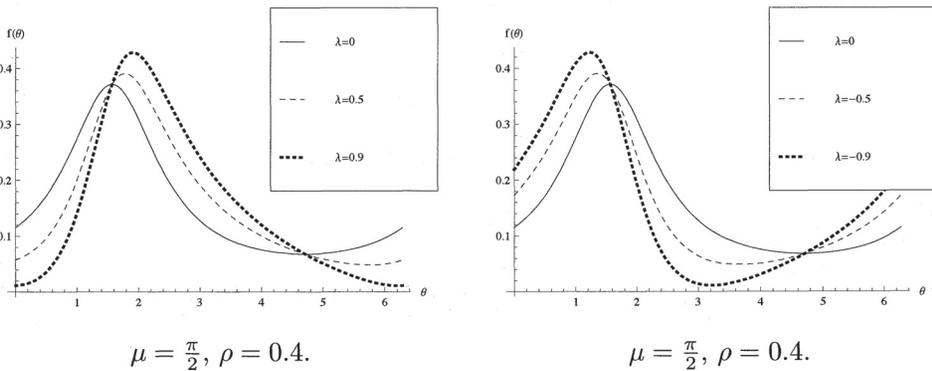


図2 巻き込みコーシー分布の正弦摂動の密度関数のプロット

コーシー分布の正弦摂動の三角モーメントも次のように陽に与えられる.

$$\begin{aligned}\alpha_p^* &:= E\{\cos(p\Theta)\} = \cos(p\mu)\rho^{|p|} - \sin(p\mu)\frac{\lambda}{2}(\rho^{|p-1|} - \rho^{|p+1|}), \\ \beta_p^* &:= E\{\sin(p\Theta)\} = \sin(p\mu)\rho^{|p|} + \cos(p\mu)\frac{\lambda}{2}(\rho^{|p-1|} - \rho^{|p+1|}), \quad (p \in \mathbb{Z}).\end{aligned}$$

これより, 平均方向は $\mu = \arg\{\alpha_1^* + i\beta_1^*\}$ であり, 平均合成ベクトル長は $\rho = \sqrt{\rho^2 + \lambda^2(1 - \rho^2)^2/4}$ となる. SSWC は常に単峰である (Abe and Pewsey (2011)). このように, 正弦摂動による非対称分布は, 三角モーメントが陽に与えられるため, 次章以降で説明する, 混合分布モデルの識別可能性や最尤推定量の一致性の条件等が容易に確認できる.

Abe and Pewsey (2011) は特別な場合として, 参照となる対称分布 $f_0(\theta)$ を Jones–Pewsey 分布の確率密度関数とした sine skewed Jones–Pewsey (SSJP) 分布族を提案している. この分布族は, $4 - 3\sqrt{3} \leq \psi \leq -0.5$ ならば常に単峰となる. SSJP の確率密度関数は $4 - 3\sqrt{3} \leq \psi \leq -0.5$ よりも広い範囲で単峰になっているようであるが, これに関する必要十分条件は得られていない. SSJP で特に, $\psi = -1$ のときの sine skewed wrapped Cauchy (SSWC) 分布の mode と antimode は陽的に与えることができ, それぞれ,

$$\begin{aligned}\theta_{mode}^* &= -\arg(\lambda + ia) + \cos^{-1}\left(\frac{a\lambda}{\sqrt{\lambda^2 + a^2}}\right), \\ \theta_{antimode}^* &= \begin{cases} -\arg(\lambda + ia) - \cos^{-1}\left(\frac{a\lambda}{\sqrt{\lambda^2 + a^2}}\right), & 0 \leq \lambda \leq 1, \\ -\arg(\lambda + ia) - \cos^{-1}\left(\frac{a\lambda}{\sqrt{\lambda^2 + a^2}}\right) + 2\pi, & -1 \leq \lambda \leq 0, \end{cases}\end{aligned}$$

である. ここで, $a = 2\rho/(1 + \rho^2)$ ($\in [0, 1)$) である. 上の例のように, $f_0(\theta)$ が wrapped Cauchy 分布の確率密度関数ならば, 摂動後もパラメータの値によらず単峰となるが, SSJP 分布族のように, 一般には, $f_0(\theta)$ が単峰であっても摂動後の分布はパラメータの選び方によっては単峰とは限らない.

非対称分布を表現する二つ目の手法は, Jones and Pewsey (2012) で紹介された基準となる対称分布の尺度変換を利用した手法であり, 単峰性を持つ非対称分布を得ることができる. 基準となる対称分布の密度関数を $f_0(\theta)$ とし, 一般性を失うことなく平均方向を $\mu = 0$ とすれば, 次の尺度変換 τ によって非対称分布を得ることができる.

$$f_0(\theta) \xrightarrow{\tau} f(\tau(\theta))$$

ただし, 関数 τ は単調増加関数である. この方法は Batschelet (1981) によって提案された手法に基づいているため, 逆 Batschelet 分布と呼ばれている.

例えば $\tau(\theta) = s_\lambda^{-1}(\theta)$, $s_\lambda(\theta) = \theta - \lambda - \lambda \cos \theta$ ととり, 基準となる対称分布を von Mises 分布にすれば, 次の von Mises 分布の尺度変換の密度関数が得られる.

$$f_{IB}(\theta|\mu, \kappa, \lambda) = \frac{1}{2\pi I_0(\kappa)} \exp [\kappa \cos \{s_\lambda^{-1}(\theta - \mu)\}]. \quad (3)$$

図 3 には, 様々な λ と異なる集中度 $\kappa \in \{1, 5\}$ に対する von Mises 分布の尺度変換の密度関数をプロットした.

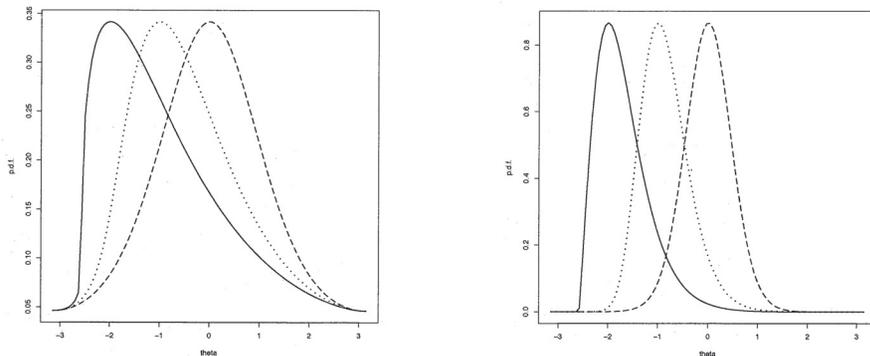


図 3 von Mises 分布の尺度変換による密度関数のプロット. $\mu = 0$, $\kappa = 1$, $\lambda \in \{0, 0.5, 1\}$ とした場合 (左), $\mu = 0$, $\kappa = 5$, $\lambda \in \{0, 0.5, 1\}$ とした場合 (右)

この分布は次のようにして密度関数であることが確認できる.

$$\begin{aligned} 1 &= \int_{-\pi}^{\pi} \frac{1}{2\pi I_0(\kappa)} \exp \{\kappa \cos \theta\} \{1 + \lambda \sin \theta\} d\theta \\ &= \int_{-\pi}^{\pi} \frac{1}{2\pi I_0(\kappa)} \frac{d(\theta - \lambda - \lambda \cos \theta)}{d\theta} \exp \{\kappa \cos \theta\} d\theta \\ &= \int_{-\pi}^{\pi} \frac{1}{2\pi I_0(\kappa)} \exp \{\kappa \cos s_\lambda^{-1}(t)\} dt. \end{aligned}$$

ここで, 変数変換 $s_\lambda(x) = x - \lambda - \lambda \cos x$ による置換積分を用いた. Jones and Pewsey (2012) の 2.2 節では “Moments of the new distributions are not especially tractable, being at essentially the same level of tractability as those of direct Batschelet distributions.” と述べているが, 上の逆 Batschelet 分布に関して言えば, p 次の三角モーメントは陽に表すことができ, さらに, この分布以外にもモーメントを求めることができる例がある. このことから, このクラスの分布族のモーメントは, 前節の Batschelet 分布族のモーメントと同程度に扱いにくい, というほど困難なものではない (例えば Abe (2015) を見よ).

角度分布の歪度の指標には、一般に平均方向 μ_1 周りの 2 次正弦モーメント $\bar{\beta}_2 \equiv E\{\sin 2(\Theta - \mu_1)\}$ と平均合成ベクトル長 ρ を使って、 $\gamma_1 = \bar{\beta}_2/(1 - \rho)^{3/2}$ で与えられる。

これより、($\mu = 0$ として) von Mises 分布の正弦振動や巻き込みコーシー分布の正弦振動の確率分布の歪度は次のように与えられる:

$$\gamma_{1,SSvM} = -\frac{2\lambda^3}{(\kappa^2 + \lambda^2)(1 - \rho_{1,SSvM})^{3/2}} \frac{I_2(\kappa)}{\kappa I_0(\kappa)},$$

$$\gamma_{1,SSWC} = -\frac{\rho\lambda(1 - \rho^2)}{2(1 - \rho_{1,SSWC})^{3/2}}$$

である。ここで、

$$\rho_{1,SSvM} = \frac{I_1(\kappa)}{\kappa I_0(\kappa)} \sqrt{\kappa^2 + \lambda^2}, \quad \mu_1 = \arg(\kappa + i\lambda),$$

$$\rho_{1,SSWC} = \sqrt{\rho^2 + \lambda^2(1 - \rho^2)^2/4}, \quad \mu_1 = \arg(\rho + i\lambda(1 - \rho^2)/2)$$

である。大きさ n の角度データの標本 $\theta_1, \theta_2, \dots, \theta_n$ が与えられたときに、標本データの対称性を検定したい。対称性の検定には、Pewsey (2002, 2004) で提案されている次の検定統計量を用いる。

$$T = \frac{n^{-1/2} \sum_{i=1}^n \sin(2(\theta_i - \bar{\theta}))}{\sqrt{\widehat{\text{Var}}[\sin(2(\theta_i - \bar{\theta}))]}}$$

ここで $\bar{\theta}$ は平均方向 μ の推定値で、 $\widehat{\text{Var}}[\sin(2(\theta_i - \bar{\theta}))]$ は $\text{Var}[\sin(2(\theta_i - \bar{\theta}))]$ の一致推定量である。検定統計量は対称性の帰無仮説の下で、漸近的に正規分布に従う。

5 非対称分布の有限混合モデル

この章では Miyata et al. (2018) に従って、以下の有限混合モデルにおける最尤推定量を求めるためのアルゴリズムを紹介する:

$$f_{msc}(\theta|\gamma) = \sum_{j=1}^g \alpha_j f_{sc}(\theta|\gamma_j), \quad (4)$$

ただし g を混合数、 $\alpha_1, \dots, \alpha_g$ を $\sum_{j=1}^g \alpha_j = 1$ を満たす混合比率、 $\gamma_j = (\mu_j, \rho_j^T, \lambda_j)^T$ とし、 j 番目の円周分布の密度関数は sine-skewed タイプの $f_{sc}(\theta|\gamma_j) = f_0(\theta|\mu_j, \rho_j)\{1 + \lambda_j \sin(\theta - \mu_j)\}$ とする。今後、 j 番目の円周分布のことを、 j 番目のコンポーネントという。尚、 $f_0(\theta|\mu_j, \rho_j)$ は任意の対称な円周分布の密度関数とし、 $\rho_j \in K \subseteq \mathbb{R}^k$, $0 \leq \mu_j < 2\pi$ を位置パラメータ、 $-1 \leq \lambda_j \leq 1$ を非対称パラメータとする。

以降、ベクトル \mathbf{a} の転置を \mathbf{a}^T と表記する. $\mathbf{Z}_i = (Z_{i1}, \dots, Z_{ig})^T$ は次のことを満たす潜在的な確率ベクトルとする:

$$\begin{aligned} \Theta_i | (\mathbf{Z}_i = \mathbf{e}_j) &\sim j \text{ 番目のコンポーネントの密度関数 } f_{sc}(\theta | \gamma_j), \\ P(\mathbf{Z}_i = \mathbf{e}_j) &= \alpha_j, \quad (j = 1, 2, \dots, g), \end{aligned}$$

ただし $\mathbf{e}_j = (0, \dots, 0, \underset{j}{1}, 0, \dots, 0)^T$ とし, $\Theta_i | (\mathbf{Z}_i = \mathbf{e}_j)$ は, $(\mathbf{Z}_i = \mathbf{e}_j)$ が与えられた下での条件付き確率変数とする. この定義から, もし Θ_i が j 番目のコンポーネントを表す母集団から生成されたのであれば, $Z_{ij} = 1$ となり, それ以外の場合には, $Z_{ij} = 0$ となることがわかる. 次に完全尤度 (complete log-likelihood) を $\ell_c(\boldsymbol{\gamma}) := \sum_{i=1}^n \sum_{j=1}^g Z_{ij} \{\log \alpha_j + \log f_{sc}(\theta_i | \gamma_j)\}$ と定義する. このとき, EM アルゴリズムは以下の形で与えられる.

初期値. 球面上の k -平均法 (Dhillon and Modha (2001)) を用いた後で, 適切な方法で初期値を表すベクトル $\boldsymbol{\gamma}^{(0)}$ を求める.

E ステップ. $\boldsymbol{\theta} = (\theta_1, \dots, \theta_n)^T$ および $(k-1)$ 回目の推定値 $\boldsymbol{\gamma}^{(k-1)}$ が与えられたもとでの完全尤度 $\ell_c(\boldsymbol{\gamma})$ の条件付き期待値を求める. これは以下の形で表される:

$$\begin{aligned} Q(\boldsymbol{\gamma}, \boldsymbol{\gamma}^{(k-1)}) &= E[\ell_c(\boldsymbol{\gamma}) | \boldsymbol{\theta}, \boldsymbol{\gamma}^{(k-1)}] \\ &= \underbrace{\sum_{j=1}^g \sum_{i=1}^n z_{ij}^{(k)} \{\log \alpha_j + \log f_{sc}(\theta_i | \gamma_j)\}}_{Q_j(\boldsymbol{\gamma}_j, \boldsymbol{\gamma}^{(k-1)})}, \end{aligned}$$

ただし $\gamma_j^{(k-1)}$ は γ_j に対する $(k-1)$ 回目の推定値とする. また

$$z_{ij}^{(k)} := P(Z_{ij} = 1 | \theta_i, \boldsymbol{\gamma}^{(k-1)}) = \frac{\alpha_j^{(k-1)} f_{sc}(\theta_i | \gamma_j^{(k-1)})}{\sum_{j=1}^g \alpha_j^{(k-1)} f_{sc}(\theta_i | \gamma_j^{(k-1)})}, \quad (j = 1, \dots, g).$$

とする.

M ステップ. $\alpha_j^{(k)} = (1/n) \sum_{i=1}^n z_{ij}^{(k)}$ とおく. されに関数 Q_j を $\lambda_j \leq 0$ の領域および $\lambda_j \geq 0$ の領域で最大化する値 γ_j^+ , γ_j^- を求める. すなわち

$$\begin{aligned} \gamma_j^+ &:= \operatorname{argmax}_{\mu_j, \rho_j, \lambda_j \geq 0} \left\{ Q_j(\boldsymbol{\gamma}_j, \boldsymbol{\gamma}^{(k-1)}) \right\}, \quad \text{and} \\ \gamma_j^- &:= \operatorname{argmax}_{\mu_j, \rho_j, \lambda_j \leq 0} \left\{ Q_j(\boldsymbol{\gamma}_j, \boldsymbol{\gamma}^{(k-1)}) \right\} \quad (j = 1, \dots, g). \end{aligned}$$

となる。このとき、 $\gamma_j^{(k)}$ と $\gamma^{(k)}$ を以下のように更新する:

$$\gamma_j^{(k)} := \begin{cases} \gamma_j^+ & \text{if } Q_j(\gamma_j^+, \gamma^{(k-1)}) \geq Q_j(\gamma_j^-, \gamma^{(k-1)}), \\ \gamma_j^- & \text{otherwise} \end{cases}$$

また $\gamma^{(k)} := (\alpha_1, \dots, \alpha_g, \gamma_1^{(k)T}, \dots, \gamma_g^{(k)T})^T$ である。

繰り返し。E ステップおよび M ステップは収束条件 (例えば、 $\epsilon > 0$ をある小さな数とし、 $\{|\ell(\gamma^{(k+1)})/\ell(\gamma^{(k)})\} - 1\} < \epsilon$ ならばアルゴリズムを終了する) を満たす、もしくは繰り返し数の上限に達するまで、E ステップと M ステップを繰り返す。

EM アルゴリズムにおける通常の M ステップとここで紹介した修正した M ステップは、いずれも関数 $Q(\bullet, \gamma^{(k-1)})$ を最大化することを目標にしている。しかしながら、真の混合数が推定に用いる混合モデルの混合数よりも小さいときには、通常の M ステップを用いた場合、EM アルゴリズムが局所解に陥ってしまう可能性がある。修正 M ステップは局所解の問題を緩和するために作られている。

6 最尤推定量の強一貫性

この章においては、SSvM 分布および SSWC 分布など位置母数を持つ円周上の確率分布の有限混合モデルに基づく最尤推定量が強一貫性を持つための条件を与える。ここで、より一般的な有限混合モデルの確率密度関数を考える:

$$f(\theta|\gamma) = \sum_{j=1}^g \alpha_j f(\theta|\mu_j, \beta_j), \quad \theta \in [0, 2\pi), \quad (5)$$

ただし g を混合数、 $\alpha_1, \dots, \alpha_g$ を混合比率、 j 番目のコンポーネントに対応する確率密度関数を $f(\theta|\mu_j, \beta_j)$ とする。密度関数 $f(\theta|\mu_j, \beta_j)$ は位置パラメータ $\mu_j \in [0, 2\pi)$ を持ち、そのほかに形状パラメータもしくはスケールパラメータに相当するベクトル β_j も含めてよいことにする。さらに B を \mathbb{R}^b におけるある部分集合とし、 $\gamma = (\alpha_1, \dots, \alpha_g, \mu_1, \beta_1^T, \dots, \mu_g, \beta_g^T)^T$ をパラメータ空間

$$\Gamma = \left\{ \gamma \mid 0 \leq \alpha_j \leq 1, \sum_{i=1}^g \alpha_i = 1, 0 \leq \mu_j < 2\pi, \beta_j \in B, \quad (j = 1, \dots, g) \right\},$$

における要素とする。Redner (1981) のアプローチにより最尤推定量の強一貫性を示すためにはパラメータ空間がコンパクトであるという条件が必要になる。しかしパラメータ空

間 Γ は非コンパクトであるため、写像 $\boldsymbol{\mu}_j := \psi(\mu_j) = (\cos \mu_j, \sin \mu_j)^T$ を用いて、位置パラメータ $\mu_j \in [0, 2\pi)$ を持つ確率モデルを以下の形に変換する:

$$f^\psi(\theta|\boldsymbol{\eta}) = \sum_{j=1}^g \alpha_j f^\psi(\theta|\boldsymbol{\mu}_j, \boldsymbol{\beta}_j), \quad \theta \in [0, 2\pi),$$

ただし $\boldsymbol{\eta} = (\alpha_1, \dots, \alpha_g, \boldsymbol{\mu}_1^T, \boldsymbol{\beta}_1^T, \dots, \boldsymbol{\mu}_g^T, \boldsymbol{\beta}_g^T)^T$ とし、パラメータ空間を

$$\Gamma^\psi := \left\{ \boldsymbol{\eta} \mid \sum_{i=1}^g \alpha_i = 1, \alpha_j \geq 0, \|\boldsymbol{\mu}_j\| = 1, \boldsymbol{\beta}_j \in B \subseteq \mathbb{R}^b \quad (j = 1, \dots, g) \right\}$$

とする。尚、 Γ^ψ の次元は $g(b+3)$ であるが、実際のパラメータの個数は $g + gb + g - 1 = g(2+b) - 1$ であることに注意されたい。

例えば、SSWC 分布の密度関数については、以下のように変形される:

$$f_{SSWC}(\theta|\boldsymbol{\mu}, \rho, \lambda) = \frac{1 - \rho^2}{2\pi(1 + \rho^2 - 2\rho\psi(\theta)^T\psi(\boldsymbol{\mu}))} \left\{ 1 + \lambda\psi(\boldsymbol{\mu})^T \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \psi(\theta) \right\}. \quad (6)$$

このとき、 $f^\psi(\theta|\boldsymbol{\mu}_j, \boldsymbol{\beta}_j)$ に相当する密度関数は以下の形で与えられる:

$$f_{SSWC}^\psi(\theta|\boldsymbol{\mu}, \rho, \lambda) = \frac{1 - \rho^2}{2\pi(1 + \rho^2 - 2\rho\psi(\theta)^T\boldsymbol{\mu})} \left\{ 1 + \lambda\boldsymbol{\mu}^T \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \psi(\theta) \right\}.$$

また Γ から Γ^ψ への全単射である写像

$$\Psi \{(\mu_1, \boldsymbol{\beta}_1^T, \dots, \mu_g, \boldsymbol{\beta}_g^T, \boldsymbol{\alpha}^T)^T\} = (\psi(\mu_1)^T, \boldsymbol{\beta}_1^T, \dots, \psi(\mu_g)^T, \boldsymbol{\beta}_g^T, \boldsymbol{\alpha}^T)^T$$

を定義する。これを用いると、 $\boldsymbol{\eta}^0 := \Psi(\boldsymbol{\gamma}^0)$ は Γ^ψ において、 $\boldsymbol{\gamma}^0$ に対応する真のパラメータとなる。

一致性を示す前に、有限混合モデルは識別不可能なモデルであることに注意されたい。例えば、真のモデルを $f_0(\theta) = f_{SSWC}(\theta|0, 0.5, 0)$ (SSWC 分布において $\mu = 0$, $\rho = 0.5$, $\lambda = 0$ とおいた) とし、推定するために用いるモデルを $f(\theta|\boldsymbol{\gamma}) = \alpha f_{SSWC}(\theta|\mu_1, 0.5, 0) + (1 - \alpha)f_{SSWC}(\theta|0, 0.5, 0)$ とすると、真のパラメータに対応する集合 $\Gamma_0 = \{(\alpha, \mu_1) | \alpha = 0 \text{ もしくは } (0 < \alpha \leq 1 \text{ かつ } \mu_1 = 0)\}$ は 1 点の集合とはならない。

これより、真のパラメータと同一視できるパラメータの集合を以下の式で定義する:

$$\Gamma^\psi(\boldsymbol{\eta}^0) = \{\boldsymbol{\eta} \in \Gamma^\psi | f^\psi(\theta|\boldsymbol{\eta}) = f^\psi(\theta|\boldsymbol{\eta}^0)\}.$$

Ω_1 および Ω_2 を \mathbb{R}^k における任意の集合とする。ここで 2 つの集合に対する距離を以下のように定める:

$$\text{dis}(\Omega_1, \Omega_2) = \text{dis}(\Omega_2, \Omega_1) = \inf_{x \in \Omega_1} \inf_{y \in \Omega_2} \|x - y\|,$$

ただし $\|\cdot\|$ をユークリッドノルムとする. もし Ω_1 および Ω_2 が 1 点からなる集合とするときには, この距離は通常のユークリッド距離になる. 加えて, 任意の点 $\omega_1 \in \Omega_1$ において, $\text{dis}(\Omega_1, \Omega_2) \leq \text{dis}(\{\omega_1\}, \Omega_2)$ が成り立つ.

Γ^ψ における $\boldsymbol{\eta}^*$ の近傍を $U_\delta(\boldsymbol{\eta}^*) := \{\boldsymbol{\eta} \in \Gamma^\psi \mid \|\boldsymbol{\eta} - \boldsymbol{\eta}^*\| < \delta\}$ とするが, 記号を簡略化するために今後は δ を省略する. また $\boldsymbol{\gamma}_j^0 := (\boldsymbol{\mu}_j^0, \boldsymbol{\beta}_j^{0T})^T$, $\boldsymbol{\eta}_j^0 := (\boldsymbol{\psi}(\boldsymbol{\mu}_j^0)^T, \boldsymbol{\beta}_j^{0T})^T$ とする.

ここで以下の条件を仮定する:

- [A1] (コンパクト性) B は \mathbb{R}^b におけるコンパクト部分集合とする. また $\boldsymbol{\eta}^0 \in \Gamma^\psi$ とする.
- [A2] (有界性) 任意の $\boldsymbol{\eta}_j \in S^1 \times B$ に対して, $E_{\boldsymbol{\eta}_j^0} [\log \{\max(f^\psi(\Theta|\boldsymbol{\eta}_j), 1)\}] < \infty$ とする, ただし $S^1 := \{\boldsymbol{\mu} \in \mathbb{R}^2 \mid \|\boldsymbol{\mu}\| = 1\}$, および $E_{\boldsymbol{\eta}_j^0} [g(\Theta)] := \int_0^{2\pi} g(\Theta) f^\psi(\theta|\boldsymbol{\eta}_j^0) d\theta$ とする. 任意の $\boldsymbol{\eta}_j^* \in S^1 \times B$ に対して, $\boldsymbol{\eta}_j^*$ の近傍 $U(\boldsymbol{\eta}_j^*)$ が存在して, $E_{\boldsymbol{\eta}_j^0} [\log \sup_{\boldsymbol{\eta}_j \in U(\boldsymbol{\eta}_j^*)} \{\max(f^\psi(\Theta|\boldsymbol{\eta}_j), 1)\}] < \infty$ となる.
- [A3] (下に有界) 任意の $\boldsymbol{\eta}_j \in S^1 \times B$ に対して, $E_{\boldsymbol{\eta}_j^0} \{\log f^\psi(\Theta|\boldsymbol{\eta}_j)\} > -\infty$ が成り立つ.
- [A4] (連続性) 任意の θ に対して, $f^\psi(\theta|\boldsymbol{\eta}_j)$ は連続である, 即ち $\lim_{\boldsymbol{\eta}_j \rightarrow \boldsymbol{\eta}_j^*} f^\psi(\theta|\boldsymbol{\eta}_j) = f^\psi(\theta|\boldsymbol{\eta}_j^*)$.

定理 1 観測値 $\Theta_1, \dots, \Theta_n$ は互いに独立に確率密度関数 $f^\psi(\theta|\boldsymbol{\eta}^0)$ に従う確率変数とする. また $\hat{\boldsymbol{\eta}}_n$ は最尤推定量を表す集合 $\text{argmax}_{\boldsymbol{\eta} \in \Gamma^\psi} \sum_{i=1}^n \log f^\psi(\theta_i|\boldsymbol{\eta})$ の任意の要素とする. このとき, 仮定 [A1]–[A4] の下で, 以下のことが成り立つ:

$$P_{\boldsymbol{\gamma}^0} \left[\lim_{n \rightarrow \infty} \text{dis} \{ \hat{\boldsymbol{\eta}}_n, \Gamma^\psi(\boldsymbol{\eta}^0) \} = 0 \right] = 1. \quad (7)$$

ただし式 (7) が成り立つならば, 自動的に $\text{dis} \{ \Gamma^\psi(\hat{\boldsymbol{\eta}}_n), \Gamma^\psi(\boldsymbol{\eta}^0) \} \rightarrow 0$ a.s. $P_{\boldsymbol{\gamma}^0}$ が成り立つことに注意されたい. 尚, 定理 1, 系 2, 命題 3 の証明については, Miyata et al. (2018) で与えられている.

系 2 観測値 $\Theta_1, \dots, \Theta_n$ は互いに独立に確率密度関数 $f(\theta|\boldsymbol{\gamma}^0)$ に従う確率変数とする. また $\hat{\boldsymbol{\gamma}}_{ML}$ は最尤推定量を表す集合 $\text{argmax}_{\boldsymbol{\gamma} \in \Gamma} \sum_{i=1}^n \log f(\theta_i|\boldsymbol{\gamma})$ の任意の要素とする. このとき, 仮定 [A1]–[A4] の下で, 以下のことが成り立つ:

$$P_{\boldsymbol{\gamma}^0} \left[\lim_{n \rightarrow \infty} \text{dis} \{ \Psi(\hat{\boldsymbol{\gamma}}_{ML}), \Psi(\Gamma(\boldsymbol{\gamma}^0)) \} = 0 \right] = 1, \quad (8)$$

ただし $\Psi(A)$ は集合 A の像とする. さらに, 任意の $\theta \in [0, 2\pi)$ に対して, $f^\psi(\theta|\boldsymbol{\eta})$ は $\boldsymbol{\eta}$ に

関してリプシッツ条件を満たすと仮定すると,

$$P_{\gamma^0} \left(\lim_{n \rightarrow \infty} \sup_{\theta \in [0, 2\pi)} |f(\theta|\hat{\gamma}_{ML}) - f(\theta|\gamma^0)| = 0 \right) = 1$$

が成り立つ.

次に系 2 を SSWC 分布の有限混合モデルに対して適用する

$$f_{MSSWC}(\theta|\gamma) = \sum_{j=1}^g \alpha_j f_{SSWC}(\theta|\mu_j, \rho_j, \lambda_j), \quad (9)$$

ただし g は混合数を表し, $f_{SSWC}(\theta|\mu_j, \rho_j, \lambda_j)$ は式 (2) で定義された密度関数とし, $\gamma = (\alpha_1, \dots, \alpha_g, \mu_1, \dots, \mu_g, \rho_1, \dots, \rho_g, \lambda_1, \dots, \lambda_g)^T$ はパラメータとする.

命題 3 観測される確率変数 $\Theta_1, \Theta_2, \dots, \Theta_n$ は互いに独立にモデル (9) に従うとし, 真値 γ_0 は制約付きパラメータ空間

$$\Gamma_c = \left\{ \gamma \mid \sum_{i=1}^g \alpha_i = 1, \alpha_j \geq 0, 0 \leq \mu_j < 2\pi, c \leq \rho_j \leq 1 - c, -1 \leq \lambda_j \leq 1 \quad (j = 1, \dots, g) \right\}$$

に属するものとする. ただし $c > 0$ はある定数とする. $\hat{\gamma}_{ML}$ を系 2 において, f を f_{MSSWC} に置き換えることでえられる最尤推定量とすると, 以下の式が成り立つ:

$$P_{\gamma^0} \left[\lim_{n \rightarrow \infty} \text{dis}\{\Psi(\hat{\gamma}_{ML}), \Psi(\Gamma_c(\gamma_0))\} = 0 \right] = 1,$$

ただし $\Gamma_c(\gamma_0) = \{\gamma \in \Gamma_c \mid f(\theta|\gamma) = f(\theta|\gamma_0)\}$ とする.

7 数値例

ここでは非対称パラメータがどのように作用するかを確認するために, 簡単な数値例を示したい. 今, 以下のパラメータを持つ 2 コンポーネントの SSWC 分布の有限混合モデルから大きさ 150 の乱数を生成した*1.

$$\begin{aligned} \gamma^0 &= (\alpha_1^0, \alpha_2^0, \mu_1^0, \rho_1^0, \lambda_1^0, \mu_2^0, \rho_2^0, \lambda_2^0)^T \\ &= (0.4, 0.6, \pi/2, 0.7, 0.5, 3\pi/2, 0.6, 0.95)^T, \end{aligned}$$

*1 生成したデータは筆者のホームページ <http://www1.tcue.ac.jp/home1/y Miyatagbt/kenkyu.html> から見る事ができる

ただし、SSWC 分布からの乱数生成は Ley and Verdebout (2017b) の p.25 に書いてある方法に従っている。その生成した乱数に対して、以下の 2 コンポーネントの有限混合モデルを用いて最尤推定量を求める。

- (1) 2つの WC 分布
- (2) 1つは WC 分布, 1つは SSWC 分布
- (3) 2つの SSWC 分布

図 4 は、上記の (1)~(3) を用いて推定した密度関数を描いたものである。尚、破線 - - - がモデル (1)、実線 — がモデル (2)、点線 がモデル (3) に対応している。特に横軸が 2.5 から 4 付近、および 5.5 から 0.5 付近の裾の振る舞いは大きく異なっていることがわかる。またモデルを評価するために赤池情報量規準 (AIC) を計算したところ、(1) の $AIC = 507.882$ 、(2) の $AIC = 504.973$ 、(3) の $AIC = 506.400$ をえた。これよりモデル選択の観点からは、モデル (2) が最も良いということになる*2。

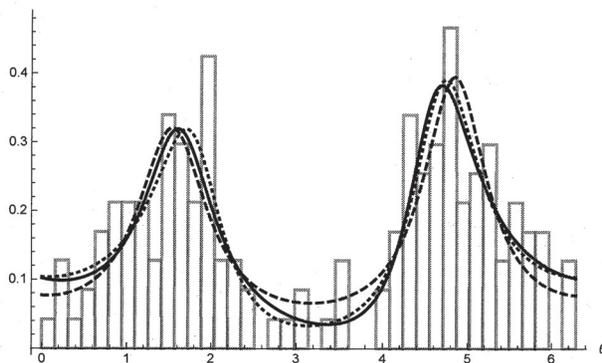


図 4 3 種類の有限混合モデル

8 まとめ

本稿では、円周上の非対称分布、およびその有限混合モデルを紹介した。さらに有限混合モデルにおける最尤推定量を求めるための EM アルゴリズム、およびその一致性の結果を Miyata et al. (2018) に従って紹介した。尚、ここで紹介した EM アルゴリズムは安定し

*2 コンポーネントが 1 つ、および 3 つの有限混合モデルにおいても AIC を求めたが、中でもモデル (2) が最も良いモデルとなった。

ているが、Mステップの最適化に多くの計算時間を必要とするため、クロスバリデーションのようなモデル評価を行うためには、不向きである。これを改善するためには、並列計算(例えば, Iida et al. (2017)), EM アルゴリズムにおける加速法(例えば, Louis (1982)), ECM(expectation-conditional maximization) アルゴリズム(例えば, Meng and Rubin (1993)) など多くの提案があるが、これらの手法の評価と有効性に関しては今後の課題とする。

謝辞

本研究は、高崎経済大学競争的研究費及び JSPS 科研費 18K01706 の助成を受けた。

参考文献

- Abe, T. (2015). Discussion: On families of distributions with shape parameters, *International Statistical Review*, **83**(2), 193-197.
- Abe, T. and Pewsey, A. (2011). Sine-skewed circular distributions, *Statistical Papers*, **52**, 683 – 707.
- Abe, T., Pewsey, A. and Shimizu, K. (2013). Extending circular distributions through transformation of argument, *Annals of the Institute of Statistical Mathematics*, **65**(5), 833–858.
- Abe, T., Shimizu, K. and Pewsey, A. (2010). Symmetric unimodal models for directional data motivated by inverse stereographic projection, *Journal of the Japan Statistical Society*, **40**, 45–61.
- Azzalini, A. and Capitanio, A. (2003). Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t -distribution, *Journal of the Royal Statistical Society, Series B*, **65**(2), 367–389.
- Batschelet, E. (1981). *Circular Statistics in Biology*. Academic Press, London.
- Banerjee, A., Dhillon, I. S., Ghosh, J. P. and Sra, S. (2005). Clustering on the unit hypersphere using von Mises-Fisher distributions, *Journal of Machine Learning Research*, **6**, 1345–1382.
- Cartwright, D. E. (1963). The use of directional spectra in studying the output of a wave recorder on a moving ship. In *Ocean Wave Spectra*, Prentice Hall, New Jersey, 203–218.

- Dhillon, I. S., and Modha, D. S. (2001). Concept decompositions for large sparse text data using clustering, *Machine Learning*, **42**, 143–175.
- Fraser, M. D., Hsu, Y.-S. and Walker, J. J. (1981). Identifiability of finite mixtures of von Mises distributions, *The Annals of Statistics*, **9**, 1130–1131.
- Holzmann, H., Munk, A. and Stratmann, B. (2004). Identifiability of finite mixtures with applications to circular distributions, *Sankhyā*, **66**, 440–449.
- Iida, M., Miyata, Y., and Shiohama, T. (2017). Bootstrap estimation and model selection for multivariate normal mixtures using parallel computing with graphics processing units, *Communications in Statistics-Simulation and Computation*, Advanced online publication: DOI:10.1080/03610918.2017.1311916.
- Jammalamadaka, S. R. and SenGupta, A. (2001). *Topics in Circular Statistics*. World Scientific Press, Singapore.
- Jeffreys, H. (1948). *Theory of Probability*, 2nd ed. Oxford University Press, Oxford.
- Jones, M. C. and Pewsey, A. (2005). A family of symmetric distributions on the circle, *Journal of the American Statistical Association*, **100**, 1422–1428.
- Jones, M. C. and Pewsey, A. (2012). Inverse Batschelet distributions for circular data, *Biometrics*, **68**, 183–193.
- Ley, C. and Verdebout, T. (2017a). Skew-rotationally-symmetric distributions and related efficient inferential procedures, *Journal of Multivariate Analysis*, **59**, 67–81.
- Ley, C. and Verdebout, T. (2017b). *Modern Directional Statistics*. CRC Press.
- Louis, T. (1982). Finding the observed information matrix when using the EM algorithm, *Journal of the Royal Statistical Society. Series B (Methodological)*, **44**(2), 226–233.
- Mardia, K. V. and Jupp, P. (2000). *Directional Statistics*. John Wiley and Sons Ltd., 2nd edition.
- Meng, X. L., and Rubin, D. B. (1993). Maximum likelihood estimation via the ECM algorithm: A general framework, *Biometrika*, **80**(2), 267–278.
- Miyata, Y., Shiohama, T., and Abe, T. (2018). Maximum likelihood estimation for finite mixtures of skew-symmetric circular distributions (投稿中).
- Pewsey, A. (2002). Testing circular symmetry, *Canadian Journal of Statistics*, **30**, 591–600.
- Pewsey, A. (2004). Testing for circular reflective symmetry about a known median axis, *Journal of Applied Statistics*, **31**, 575–585.

- Redner, R. (1981). Note on the consistency of the maximum likelihood estimate for nonidentifiable distributions, *The Annals of Statistics*, **9**, 225–228.
- 清水 邦夫 (2008). 方向統計学の最近の発展, *計算機統計学*, **19(2)**, 127–150.
- Shimizu, K. and Iida, K. (2002). Pearson type VII distributions on spheres, *Communications in Statistics – Theory and Methods*, **31**, 513–526.
- Umbach, D. and Jammalamadaka, S. R. (2009). Building asymmetry into circular distributions, *Statistics & Probability Letters*, **79**, 659–663.
- von Mises, R. (1918). Über die “Ganzzahligkeit” der Atomgewichte und verwandte Fragen, *Physikal. Z.*, **19**, 490–500.