Vol.2021-MUS-132 No.16
2021/9/17

統計学習を介する文化進化のモデルと音楽・文芸・絵画データにおける共役分布則

中村 栄太[1,a)]　持橋 大地[2]　齋藤 康之[3]

概要：本発表では，音楽・文芸・絵画データで見つかった分布則とその起源が説明可能な文化進化モデルについて議論する．近年の研究により，音楽創作などの知的行動を特徴付ける多くの文化形質は，統計学習を介して世代間で伝達されることが示唆されている．そこで，統計学習が文化進化に与える影響について考える．まず，大規模な音楽データ解析によって，音楽スタイルを表す様々な形質のデータセット内での分布がベータ分布やその他の共役分布に従うことを観察する．次に，統計学習を含む文化進化の一般的なモデルを構成して，斜行伝達が存在する場合に共役分布が平衡状態で現れることを解析的に示す．この結果は文化生産物の生成過程を表す確率分布と対応する文化形質の集団内分布の形を定量的に関係付ける点で興味深い．この理論の応用として，少量のデータからの形質分布の効率的な推定法についても議論する．さらに，文芸データと絵画データにおいて同様の分布則が見られることを確認する．

1. はじめに

文化の伝達と進化は人間社会の発展にとって本質的である [1]．これらの過程を調べるため，文化形質の伝達，選択，変異の過程を取り入れた力学系あるいは確率過程を解析する研究が行われている [2,3]．例えば個人の母国語（英語かフランス語かなど）のような文化形質に対しては，すでに発展している遺伝子や生態系の進化の理論を適用できる [2,4,5]．しかし，例えば音声生成のスタイル（アクセントや話速など）や音楽の創作スタイルなどの知的行為を特徴付ける多くの文化形質は，遺伝子の複製などとは大きく異なる方法で伝達されている．情報学や認知科学の分野では，こうした知的行為は確率的データ生成過程を含み，その過程を制御する統計的パラメーターは文化生産物の統計学習を通して世代を超えて伝達されることが示唆されている（例えば [6–8]）．そこで，文化進化と社会的学習において統計学習を含む伝達過程の兆候や帰結を調べる研究が盛んである [9–14]．

複雑な生物システムや社会システムにおいて，背後にある動的過程の痕跡はしばしば系の集団的振る舞いを特徴付ける分布則として見出される．それらの分布則は，見つかれば，モデル構築における制約を与え，現象の理論的理解を深める点で重要である．そうした分布則の例として，相互依存する個体の移動ダイナミクス [15] や自己複製する細胞内の化学反応ダイナミクス [16] に見られるジップ則や，交通網のランダム拡散過程に見られるテイラー則 [17] などがよく知られている．近年研究が活発化している文化生産物の大規模データ解析 [18–20] では，同様の分布則として，西洋クラシック音楽における特定の音楽要素の曲中頻度が近似的にベータ分布に従っているとの報告がある [21]．ベータ分布はベイズ統計において共役分布の例として知られている．共役分布とは，ある統計モデルに対して事前分布と事後分布の両方をパラメータ化できる分布である [22]．この観察は文化形質の分布形とその形質の伝達に関わる統計学習過程との間の関係を示唆しているが，この関係性を理解する理論的枠組みはこれまで確立されていない．

本研究の目的は，統計学習を介する伝達過程によって文化形質の分布がどう時間変化するかを定量的に理解することである．まず，音楽データにおいて様々な統計的形質を分析して経験的な分布則を抽出する．異なる社会で作られた音楽データの様々な頻度統計量の分析に加えて，音符密度や音程スケールなどの全体的な音楽統計量の分析も行う．これにより形質分布と文化生産の統計モデルとの間の一般的関係が示唆されることを見る．次に，統計学習を組み込んだ文化進化のモデルを調べて，伝達過程に関する一般的な仮定の下でデータ生産モデルと対応する形質の平衡分布との間の関係性が導かれることを見る．この理論的結果は経験的な分布則を説明するとともにそれを広いクラスの文化データ生産過程に一般化するものである．また，理

1 京都大学 Kyoto 606-8501, Japan
2 統計数理研究所 Tokyo 190-8562, Japan
3 木更津工業高等専門学校 Chiba 292-0041, Japan
a) eita.nakamura@i.kyoto-u.ac.jp

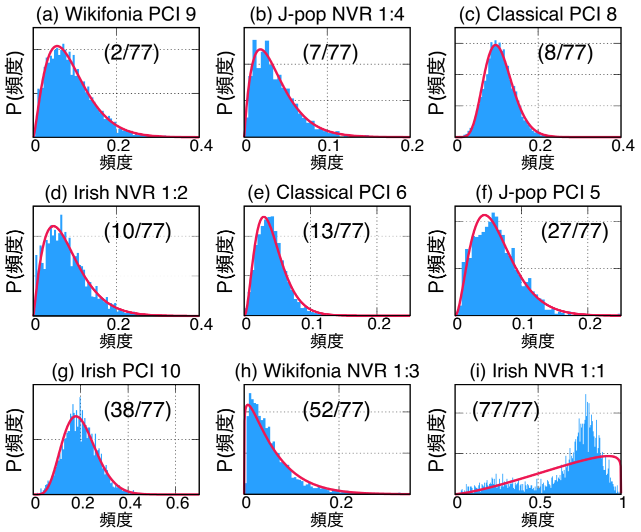

図 1 音高クラス音程（PCI）と音価比（NVR）の曲中頻度のデータセット内分布とそれをベータ分布でフィットしたもの．“PCI i” は i 半音の PCI を表す．パネル内の数字はフィットの良さのランクを示す（フィットの良さはデータ分布とフィットしたベータ分布との差の累積値で計算している）．

論の応用として，少量データからの形質分布の効率的な推定法について議論する．さらに，文芸と絵画データにおいて同様の分布則を検証して，分析した形質の多くで理論から期待される共役分布が見られることを確認する．

2. 音楽データにおける経験的分布則

文化データにおける経験的分布則を抽出して，特にベータ分布の普遍性を検証するため，異なる社会で作られた音楽の 4 つのデータセットを分析して，音高とリズム要素の頻度の分布を観測する．これらのデータ内の曲は，MIDI あるいは MusicXML 形式の楽譜データとして与えられており，各音符の音高は半音単位の整数値で表される．MusicXML ファイルでは，音の長さ（音価）と小節線の情報も取得でき，音符系列を小節ごとに区切ることも可能である．「クラシック音楽」データ [21] には，多様な楽器編成の曲が含まれており，音高情報を取得できる．他のデータセットは全てメロディーのデータであり，音高と音価の情報を取得できる．「Wikifonia」データには主に 20 世紀前半から中盤にかけて作られた西洋のポピュラー音楽およびジャズ音楽の曲が含まれている [23]．「J ポップ」データには，1950 年以降に作られた日本のポピュラー音楽の曲が含まれている [24]．「アイリッシュ音楽」データには，Web 上の公開サイトから集取したアイリッシュ音楽のメロディーが含まれている [25]．これらのデータセットには，どれも様々な作曲家の作った曲が，それぞれ計 $O(10^3$–$10^4)$ 音符分含まれている．文化形質としては，隣り合う音符間の音程を 12 を除数として計算した音高クラス音程（PCI; pitch-class interval）の曲中頻度および隣り合う音符の音価の比の曲中頻度を抽出して，各データセット内の分布を観察する．

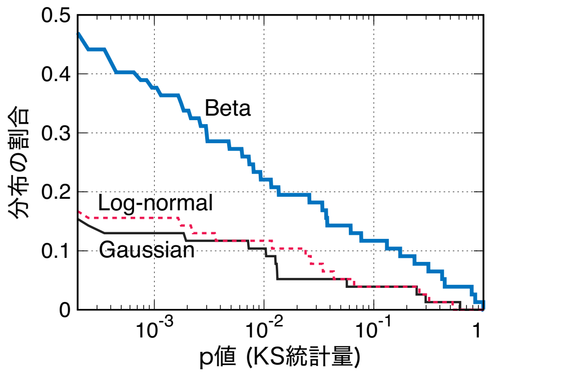

図 2 KS 統計量の p 値の相補累積分布関数．ベータ分布，対数正規分布，ガウス分布によるフィットに対する結果．

図 1 に示す通り，異なるデータセットにおいて様々な音楽要素の分布で近似的なベータ分布が見られた．これは，クラシック音楽で 2 つの要素に関して見つかったベータ分布則を拡張した結果である [21]．ベータ分布のパラメーターは 2 つのみであり，例えばこれらにより平均値と分散を調整した場合，3 次モーメントあるいは歪度は勝手に調整することはできない．図 1(i) の例が示すように，ベータ分布から大きくずれた分布は広がりを持ち，複数のピークを持つものが多かった．これらの複数ピークは，例えば 2 分リズムと付点リズムのように異なる音楽スタイルに対応することがしばしば観察された．

より形式的にフィットの良さを測り，その他の分布の仮説との比較を行うため，コルモゴロフ・スミルノフ（KS）統計量の p 値の分布を図 2 に示す．比較対象として，ベータ分布と同様に 2 つの調整可能なパラメーターを持つ対数正規分布と正規分布に対する結果も図に示している．結果では，ベータ分布からのずれが統計的に有意でない分布の割合は，95% 信頼区間に対して 14% と高くないことがわかる．一方で，他の分布に比較してベータ分布のフィットはかなり良いことが確かめられた．よって，データ分析により，パラメーターを 2 つ持つ単純な分布の候補の中ではベータ分布が最も良い近似になっていることが示された．

次に，同じデータセットを用いて，曲の全体的な特徴を捉える音楽統計量に関する分析も行った．図 3 は J ポップデータに対する結果を示している．第一に，隣り合う音高の差の絶対値として定義される音程 I の各曲中での分布に注目する．図 3(a) の様に，この分布の概形は，値域 $1 \leq I \leq 12$ において指数分布 $P(I) \propto e^{-\lambda I}$ で近似される（より細かく見れば不協和音程の頻度が低くなる傾向があることも分かる）．この指数分布のレートパラメーター λ は，各曲に対して $\lambda = 1/\langle I \rangle$ と推定できる（ここでは $\langle \cdot \rangle$ は曲中での平均値を示す）．この統計量は曲の音程の全体的なスケールを表している．曲ごとのレートパラメータ λ のデータセット内での分布 Ψ は，近似的にガンマ分布に従っている．

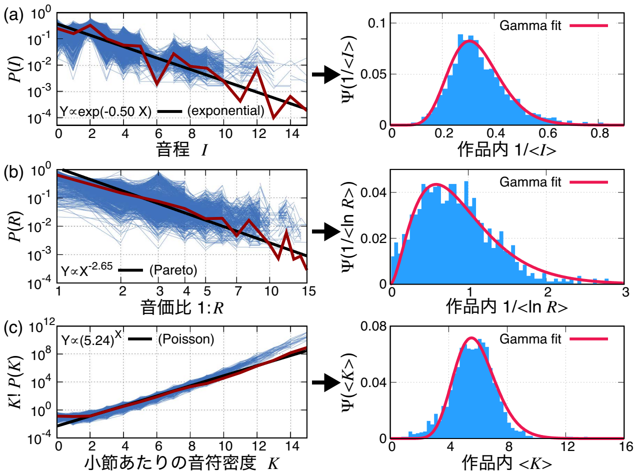

図 3　大局的音楽統計量の分布（J ポップデータ）．(a) 左：作品内の音程の分布 P（細線），データセット内での分布（太線），および指数分布によるフィット（直線）．右：音程の曲ごとの統計量（平均の逆数）の分布 Ψ とガンマ分布によるフィット．(b) $1{:}R$ の音価比に対する同様の結果．全体の分布はパレート分布でフィットしている．(c) 小節ごとの音符密度に対する同様の結果．全体の分布はポアソン分布でフィットしている．

第二に，隣り合う音価の比 $1{:}R$ の内，R が整数となるものの各曲中での分布に注目する．図 3(b) の様に，この分布の概形は，値域 $1 \leq R \leq 10$ においてパレート分布（あるいはべき乗分布）$P(R) \propto R^{-\alpha}$ で近似される．このパレート分布の形状パラメーター α は，各曲に対して $\alpha = 1/\langle \ln R \rangle$ と推定できる．この統計量は曲の中で使われる音価比の大きなリズムの全体的な量を表している．曲ごとの形状パラメーター α のデータセット内での分布 Ψ は，おおよそガンマ分布に従っている．

最後に，小節に含まれる音符数（音符密度）K の各曲中での分布に注目する．図 3(c) の様に，この分布の概形は，値域 $2 \leq K \leq 12$ においてポアソン分布 $P(K) \propto \rho^K/K!$ で近似される．このポアソン分布のレートパラメーター ρ は，各曲に対して $\rho = \langle K \rangle$ と推定できる．この統計量は曲の全体的な音符密度を表している．曲ごとのレートパラメーター ρ のデータセット内での分布 Ψ は，ここでもおおよそガンマ分布に従っている．他のデータセットに関しても同様の結果が得られた．

観察されたガンマ分布は，それぞれ指数分布，パレート分布，ポアソン分布の分析したパラメーターに対する共役分布になっている．ベータ分布則と合わせると，これらの観測結果は，文化生産物内での統計分布とその統計的文化形質の集団内での分布との間の一般的な共役関係を暗示している（ここではこれを共役分布則と呼ぶ）．この関係が様々な社会，時代における音楽データで観察されたことは，その背後に一般的な機構が存在することを示唆する．

3.　モデル

3.1　力学系の定式化

共役分布の起源の可能性を探るため，統計学習生成系と呼ばれる進化モデルを考える．各世代 t において N 人がそれぞれ文化生産物 X_n^t $(n = 1, \ldots, N)$ を生成するものとする．各生産物（例えば曲）$X_n^t = (x_{n\ell}^t)_{\ell=1}^L$ は L 個のサンプル（例えば音符）$x_{n\ell}^t$ から成り，それらは「データ生産モデル」と呼ばれる確率分布 $\phi(x; \theta_n^t)$ から独立に生成されるとする．パラメーター θ_n^t を文化形質として扱う．簡単のため，各個人により作られた全ての生産物の集合をここでは一つの生産物として扱い，全て同じ大きさ L を持つとする．全ての人の生産モデルは同じ形を持つが，そのパラメーターの値は違っても良いものとする．例えば，ϕ がベルヌーイ分布の時は，サンプル $x_{n\ell}^t$ は 0 または 1 の値を取り，θ_n^t は 1 が出る確率を表す．（通常は，1 は注目する特定の要素（例えば 6 半音の音高クラス音程など）を表し，0 はその他の場合を表す．）別の例として，ϕ がポアソン分布の場合は，サンプル $x_{n\ell}^t$ は非負の整数値を取り，θ_n^t はレートを表す．以降，個体 n とその文化形質 θ_n^t を同一視することがある．

次に文化形質が世代間で伝達される過程を考える．一般には，次世代において文化的子を生む個体についての選択過程を導入できる．ここでは，文化生産物の選択過程やその強さについてまだ不確定性が大きいこと [26, 27] から，選択圧がない単純な場合を考えて，伝達過程の効果に注目する．また，各個体 θ_n^{t+1} は主要な文化的親 θ_n^t（一次親と呼ぶ）を 1 人持つものとする．簡単のため，各親 θ_n^t は 1 人の文化的子 θ_n^{t+1} を持つものとする（集団サイズが無限大の極限ではランダム選択の場合と同じ解析結果になる）．文化の伝達では，親世代の他の個体（二次親と呼ぶ）の生産物も学習に用いられることがしばしばあり，斜行伝達と呼ばれる [2]．この場合，文化形質の学習に使われる生産物の集合（学習データ）は，一次親と二次親の生産物を合わせたものとなる．

形質 θ_n^{t+1} の各パラメーターは学習データから抽出した統計量の期待値として得られるものと仮定する（この仮定は後で正当化する）．生産物 X_n^t から得られる期待値を $\hat{\theta}_n^t$ と記すことにする．子が一次親の生産物のみを学習に使う場合（垂直伝達）は，$\theta_n^{t+1} = \hat{\theta}_n^t$ が成り立つ．二次親の生産物も学習に使う場合は，パラメーター θ_n^{t+1} は $\hat{\theta}_n^t$ と二次親の生産物の期待値 $\hat{\theta}_{n'}^t$ の重み付き和で与えられる．二次親がそれぞれ平等に影響を与えるとすると，斜行伝達の効果は以下の式で与えられる．

$$\theta_n^{t+1} = (1-u)\hat{\theta}_n^t + u\zeta, \quad \zeta = \frac{1}{N'} \sum_{n'} \hat{\theta}_{n'}^t. \qquad (1)$$

ここで，和は全ての二次親に対してとり，N' は二次親の総

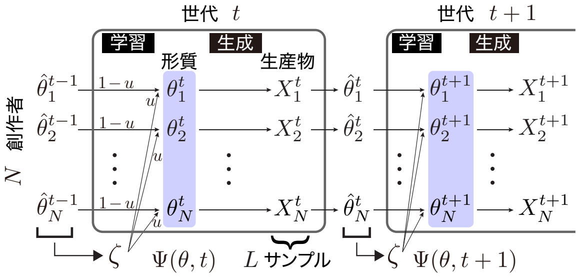

図 4　斜行伝達がある場合の統計学習生成系.

数，u はそれらの影響の大きさを表す係数である（図 4）. 以下では，二次親はランダムに選ばれ，N' が大きい時を考えるが，この場合は $\zeta \simeq \bar{\theta} = (1/N)\sum_n \theta_n^t$（集団平均）となり，$N$ が大きい時は ζ は実効的に定数として扱える.

ϕ がベルヌーイ分布に従う場合，この統計学習生成系は集団遺伝学で知られるライト・フィッシャー（WF）モデルと等価であることが次の様にわかる [28]．選択的に中立な WF モデルは N_g 個の個体からなる集団により構成され，各個体は 2 つの対立遺伝子を持つ（それぞれ 0 と 1 と表す）．各世代において，全ての個体が新しい個体に入れ換わり，ランダムに選ばれる親の遺伝子が継承される．この時，遺伝子頻度は $L = N_g$ の関係にある統計学習生成系で垂直伝達される形質と同じ確率過程に従うため，統計学習生成系は WF モデルに従う N 組の集団と等価であることがわかる．さらに，式 (1) の斜行伝達の効果は，ζ が定数である時，一定の外部からの移住あるいは突然変異を表す線形圧の効果と等価である [29].

3.2　平衡分布

文化形質の集団内分布 $\Psi(\theta,t) = P(\theta_n^t = \theta)$ の時間発展を N が大きな極限で調べる．この問題は WF モデル，あるいは ϕ がベルヌーイ分布の場合には良く理解されている [30]．純粋な垂直伝達の場合には，変数 θ_n^t は統計的に揺らぎながら最終的に 0 または 1 になった時点で変化しなくなる．この平衡分布は $\Psi(\theta,\infty) = (1-\bar{\theta})\delta(\theta) + \bar{\theta}\,\delta(\theta-1)$ で与えられる．斜行分布（あるいは線形圧）がある場合は，非自明な平衡分布が Wright により見つけられている [29]．解析解を導出するため，連続時間極限および L が大きい極限を考える．この場合のダイナミクスは次のフォッカー・プランク（FP）方程式により記述される.

$$\frac{\partial \Psi(\theta,t)}{\partial t} = -\frac{\partial}{\partial \theta}\left\{ M(\theta)\Psi(\theta,t) - \frac{1}{2}\frac{\partial [K(\theta)\Psi(\theta,t)]}{\partial \theta} \right\}. \tag{2}$$

ここで，$M(\theta) = u(\zeta - \theta)$，$K(\theta) = \theta(1-\theta)/L$ でありの右辺の中括弧の中身は確率流を表す [30]．FP 方程式のゼロ流れの平衡解はベータ分布 $\mathrm{Beta}(\theta; 2Lu\zeta, 2Lu(1-\zeta))$ により与えられることがわかる [29]．これにより，統計学習生成系は斜行伝達がある場合にベータ分布則の進化的起源を説明できる.

上の結果をより一般のデータ生産モデルに対する文化形質に拡張するため，パラメーターを θ とする次の指数型分布族を考える.

$$\phi(x;\theta) = \exp\bigl[F(x)B(\theta) - A(\theta) + U(x)\bigr]. \tag{3}$$

この形の分布は，$F(x)$ と $U(x)$ を適切に与えることで，ベルヌーイ分布，ポアソン分布，ガウス分布，ガンマ分布を含む広いクラスの確率分布を記述できる．関数 $B(\theta)$ と $A(\theta)$ は分布のパラメーター化の仕方を指定するもので，$0 = \partial_\theta \int dx\, \phi(x;\theta)$ から $\langle F \rangle = A'(\theta)/B'(\theta)$ が成り立つことがわかる．式 (3) の分布に対しては，以下のように共役分布を陽に構成することができる.

$$\tilde{\phi}(\theta;\chi,\nu) = \exp\bigl[\chi B(\theta) - \nu A(\theta) + C(\theta) + W(\chi,\nu)\bigr]. \tag{4}$$

ここで，$C(\theta)$ は任意の関数で，$W(\chi,\nu)$ は正規化項を表す．関数 C は共役性には寄与せず，ベイズ統計で通常用いられるのは $C = 0$ となる極小ケースである．パラメーター θ の座標系を固定するにあたって，期待値 $\langle F \rangle$ をパラメーターとして使うのが自然であることに注意する．これはクラメール・ラオ限界を達成できる唯一の座標系であることが知られており [31]，それは統計学習において最も効率の良いパラメーターの取り方であることを意味する．そこで，以下では $\langle F \rangle = A'(\theta)/B'(\theta) = \theta$ が成り立つと仮定する．F の推定分散は $D(\theta) \equiv \langle (F-\theta)^2 \rangle = 1/B'(\theta)$ で与えられ，分布 $\Psi(\theta,t)$ は式 (2) で $K(\theta) = D(\theta)/L$ とした方程式に従う.

斜行伝達の効果 $M(\theta) = u(\zeta - \theta)$ の下では，式 (2) のゼロ流れの平衡解 $\Psi_*(\theta)$ は次の式を満たす.

$$2Lu(\theta - \zeta)\Psi_* = -\frac{d}{d\theta}(D\Psi_*). \tag{5}$$

この方程式の解は次のように書ける.

$$\Psi_*(\theta) \propto \frac{1}{D(\theta)} \exp\left[\int^\theta \frac{2Lu(\zeta - \eta)}{D(\eta)} d\eta\right]. \tag{6}$$

$D(\theta)^{-1} = B'(\theta)$ および $\theta = A'(\theta)/B'(\theta)$ を使って，次が得られる.

$$\Psi_*(\theta) \propto \exp\bigl[2Lu\zeta B(\theta) - 2LuA(\theta) + \ln B'(\theta)\bigr]. \tag{7}$$

式 (4) と比べると，これは共役分布で $C = \ln B'$ としたものである．よって，斜行伝達がある場合に一般の共役分布則の起源が説明できることがわかる．式 (6) と (7) は集団内の文化形質の平衡分布の形が文化生産の統計モデルに依存することを示している．Ψ_* は本質的に推定分散 D により決定されることがわかる．$\ln B'(\theta)$ の項と測度 $d\theta$ の積 $B'(\theta)d\theta$ は変数変換に対して不変な測度であるため，式 (7) は変数変換に対して不変な確率分布になっている.

さらに，$\ln B'(\theta)$ が定数項を除いて $B(\theta)$ と $A(\theta)$ の線形結合となる場合に，解 (7) は極小共役分布となることもわかる.

生産モデル ϕ	文化形質 θ	形質分布 Ψ
ベルヌーイ	平均	ベータ
ポアソン	平均	ガンマ
ガウス	平均	ガウス
ガウス	分散	逆ガンマ
ガンマ	平均	逆ガンマ

表 1 式 (8) を満たす指数型分布の例と対応する共役分布．

つまり，ある定数 c_i に対して $\ln B'(\theta) = c_0 + c_1 B(\theta) + c_2 A(\theta)$ が成り立つ場合である．両辺を微分することにより等価な条件 $B''(\theta)/B'(\theta) = (c_1 + c_2\theta)B'(\theta)$ が得られる（$A'(\theta)/B'(\theta) = \theta$ を使った）．$D(\theta) = 1/B'(\theta)$ が成り立つため，これは $D'(\theta)$ が θ について線形，つまり $D(\theta)$ が θ についての 2 次式であることと等価である．以上により，平衡分布が極小共役分布となる必要十分条件は推定分散 $D(\theta) = 1/B'(\theta)$ が以下の 2 次式で与えられることである．

$$D(\theta) = d_0 + d_1\theta + d_2\theta^2. \tag{8}$$

ただし，d_i は定数を表す．この結果は，式 (4) で $C = 0$ の場合をゼロ確率流方程式に代入して，以下の解を明示的に得ることでも確認できる．

$$\chi = 2Lu\zeta - d_1, \quad \nu = 2Lu + 2d_2. \tag{9}$$

式 (8) を満たす確率分布の例を表 1 に示す．一般に $\phi(x;\theta)$ が式 (8) を満たす時，変換関数 f により得られる変数 $y = f(x)$ の分布も同じく満たすことがわかる（この時 D は変化しない）．例えば，ガンマ分布は平均統計量に関して式 (8) を満たすため，パレート分布は対数平均統計量に関して式 (8) を満たす．2 次式の推定分散を持つ指数型分布族はよく知られる確率分布の多くのものを含むことが知られており [32]，それらは図 3 で分析した分布も含む．これは文化生産物で見られる共役分布はしばしば極小型になり得ることを示している．一方で，全ての指数型分布族が式 (8) を満たす訳ではない（ガンマ分布が対数平均統計量に関して反例を与える）．

3.3 平衡分布の性質

平衡分布に関するいくつかの性質をさらなる解析により理解することができる．まず，多くの場合に満たされる正規化条件の下で，関係式 $\langle\theta\rangle_* \equiv \int d\theta\,\theta\Psi_*(\theta) = \zeta$ が成り立つ（ζ が外部パラメーターであっても成り立つ）．これは集団平均 $\zeta = \bar{\theta}$ が一定とする仮定と整合する．次に，条件式 (8) が満たされる時，次の関係式を示せる．

$$V_*(\theta) = \int d\theta\,(\theta - \langle\theta\rangle_*)^2\Psi_*(\theta) = \frac{D(\zeta)}{2Lu - d_2}. \tag{10}$$

これは統計学習生成系の個体が同時に複数の文化形質を伝達する時，平衡状態における $V(\theta)/D(\bar{\theta})$ の値は $\bar{\theta}$ に依存せず，Lu と d_2 にのみ依存することを表している．

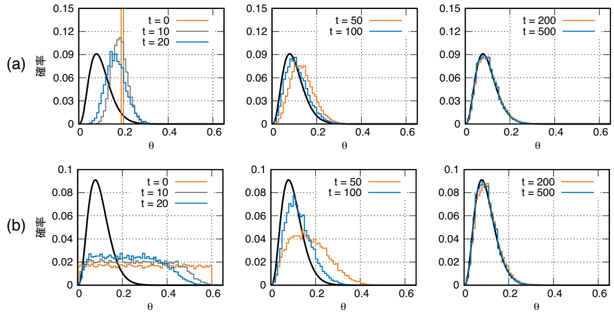

図 5 ベルヌーイ分布に従う生産モデルの統計学習生成系に対するシミュレーション結果．パラメーターは $N = 20000$, $L = 1000$, $u = 1/50 = 0.02$, $\zeta = 0.1$ を用いた．$\mathrm{Beta}(\theta; 4, 36)$ をフィットしている．(a) 初期分布が $\theta = 0.2$ に集中している場合．(b) 初期分布が区間 $[0, 0.6]$ で一様に分布している場合．

また，緩和時間のオーダーは u^{-1} で与えられる．これは斜行伝達により，単位時間当たり割合 u の形質が ζ に置き換えられることから物理的に期待される結果である．表 1 の中の分布に対しては，時間依存 FP 方程式 (2) を固有関数法を用いて解くことで，この緩和時間を陽に求めることもできる．この結果は図 5 において数値計算によっても確かめられている．この結果はベルヌーイ分布の生産モデルを持つ統計学習生成系に対してシミュレーションを行って得られたものである（ζ は定数とした）．二通りの極端な初期分布に対して，緩和時間がオーダー u^{-1} であることが見て取れる．重要なことに，この結果は，選択圧がある場合にもその効果がオーダー u^{-1} の時間スケールで小さいならば，共役分布が維持されることを示している．

u の値が実験的に求まらない限り（現在のところそれは難しい），平衡に達する実際の時間を推定することは難しい．一方で，ある社会の文化（例えば西洋クラシック音楽）で一度（準）平衡分布が達成されると，その社会での後続の文化（例えば西洋ポピュラー音楽）では，多くの伝統が受け継がれる限り，平衡状態に速く達することが期待される．例えば，世界中の様々な音楽には共通する特徴が存在することが知られており [33]，これは音楽の多くの側面が長い時間受け継がれることを示唆している．

理論を実験データと関係付けるには，創作者の形質 θ が実際には生産物を通してのみ観測できることに注意する必要がある．θ が L_p 個のサンプルからなる生産物から推定されるときには，$D(\theta)/L_p$ の分散が期待され，（図 1 や 3 で見たような）生産物ごとの統計パラメーターの分布 Ψ_p は θ の分布から変化したものになる．それでもなお，この変形の過程は伝達過程と同じであるため，Ψ_p は Ψ_* と同じ共役分布に近似的に従い，分散は式 (10) から $D(\zeta)/L_p$ だけ増加することが期待される．これにより図 1 や 3 で観察された経験的な共役分布則が説明される．この議論は同時に，$Lu \gg L_p$ となる状況では，生産物データから直接 Lu の値を推定するのに式 (10) を適用するのが難しいことも

示している．図 1 の分布から得られた $D(\bar{\theta})/V(\theta)$ の値は，Wikifonia データ（ベータ分布のフィットが著しく悪いものは除いている）に対して 29.5(平均) ± 9.11(標準偏差) であり，これは異なる音楽要素間では比較的一致しており，平均値 $\overline{L}_p = 107$ より有意に小さい．非常に小さな u の値は現実的ではないと考えられるため，これらの小さな値は音楽で頻繁に現れる反復によるサンプル間の統計的依存性を反映した結果だと考えられる．他のデータセットに対しても同様の結果が得られている．

3.4 多次元パラメーターの場合

前節までの理論解析は，データ生産モデルが多次元パラメーターを持つ指数型分布の場合に拡張でき，式 (7) を一般化した共役分布が同様に得られることが示せる．ここでは多次元パラメーターの場合の主な結果と具体例について説明する．導出は文献 [34] を参照されたい．

多次元パラメーター $\theta = (\theta_i)_{i=1}^K$ を持つ指数型分布は次の様に表される．

$$\phi(x;\theta) = \exp\left[\sum_{i=1}^{K} F_i(x)B_i(\theta) - A(\theta) + U(x)\right]. \quad (11)$$

但し，$x = (x_k)_{k=1}^{K'}$ は確率変数，F_i は独立な統計量である．$\partial_i = \partial/\partial\theta_i$ と書くと，$\partial_i A = \langle F_k\rangle\partial_i B_k = G_{ik}\langle F_k\rangle$ が成り立つ．ここで $G_{ik} \equiv \partial_i B_k$ で定義される行列は，逆行列を持つ（即ち無駄な変数はない）と仮定する．以下では，$\langle F_k\rangle = \sum_i G^{-1}_{ki}A_i = \theta_k$ で定義される期待値座標系で考える．

FP 方程式 (2) を多次元に拡張すると，右辺は各次元の確率流のダイバージェンスとして表される．詳細は省くが，ドリフト項 M はベクトルに拡張され $M_i(\theta) = u(\zeta_i - \theta_i)$ で与えられ，拡散係数 K は行列に拡張され，$K_{ij} = G^{-1}_{ij}/L$ で与えられる．この方程式の平衡解は式 (7) の一般化である次の共役分布で与えられる．

$$\Psi_* \propto \exp\left[2Lu\sum_i \zeta_i B_i(\theta) - 2LuA(\theta) + \ln\det G(\theta)\right]. \quad (12)$$

1 次元の場合と同様，最後の項は $|\partial B/\partial\theta|d\theta$ の形で不変測度を与えるため，この式は $G_{ik} = \partial_i B_k(\theta)$ と書いた時に変数変換に対して不変な確率分布になっている．

具体例として，生産モデルが離散分布である場合を考える．この場合のパラメーターは足して 1 となる確率ベクトルで与えられ，$\theta = (\theta_i)_{i=1}^{K+1}$ と書く．ここで，$K+1$ 次元の離散分布を考えるが，実質的に K 個の独立なパラメーターを持つ．$\theta_{K+1} = 1 - \sum_{i=1}^K \theta_i$ と書くと，θ_i は期待値座標系を与え，平衡分布は次のディリクレ分布となることが確認できる．

$$\Psi_* \propto \prod_{i=1}^{K+1} \theta_i^{2Lu\zeta_i - 1}. \quad (13)$$

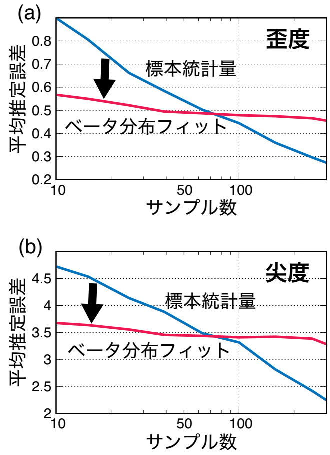

図 6 歪度と尖度の推定誤差．

もう一つの例として，生産モデルが平均と分散をパラメーターとするガウス分布の場合を考える．

$$\phi(x;\mu,\Sigma) = \mathrm{Gauss}(x;\mu,\Sigma) = \frac{1}{\sqrt{2\pi\Sigma}}e^{-(x-\mu)^2/2\Sigma}. \quad (14)$$

平均と分散のペア (μ,Σ) はこの分布の期待値パラメーターではないが，通常用いられるものであるため，ここでは平衡解の共役分布をこのパラメーターの分布として計算する．式 (12) は座標系に依らないため，直接解を計算できる．統計量を $F_1 = x$ と $F_2 = x^2$ とすると，$B_1 = \mu/\Sigma$, $B_2 = -1/2\Sigma$，および $A = \mu^2/2\Sigma + \ln(2\pi\Sigma)/2$ であり，計算により $\det G = 1/2\Sigma^3$ が得られる．これらを式 (12) に代入してまとめると，以下の共役分布が得られる．

$$\Psi_*(\mu,\Sigma) = \mathrm{Gauss}\big(\mu;\zeta_1,\Sigma/(2Lu)\big) \cdot \mathrm{InvGam}\big(\Sigma, Lu + \tfrac{3}{2}, Lu(\zeta_2 - \zeta_1^2)\big). \quad (15)$$

これはガウス・逆ガンマ共役分布としてベイズ統計で良く知られる分布形である．

4. 形質分布の推定への応用

前節の理論の応用として，共役分布則が少量データからの形質分布の効率的な推定に使えることを説明する．形質分布の詳細構造は進化のモデルの構築や検証のための重要な情報となるが，解析に使える文化生産物の数が限られている時にはこれは難しい問題となる．そのような場合には，形質分布の近似形に関する事前知識が有用な案内となる．それは，事前知識によりデータから推定すべきパラメーターの数が実効的に減らすことができるからである．

具体的な問題設定として，図 1 で分析した 77 の頻度統計量のデータを使って，各分布の歪度と尖度を推定することを考える．これらは，平均と分散に次いで，確率分布の形状を捉える最も重要な量である．形質分布に関する事前知識が無い場合には，これらの 4 つの量（平均，分散，歪度，尖度）の間にいかなる関係も仮定できないため，それらを

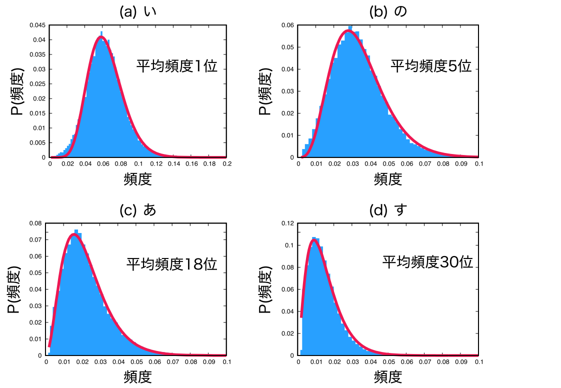

図 7 歌詞データにおける読みがなの作品内頻度の分布とベータ分布によるフィット．

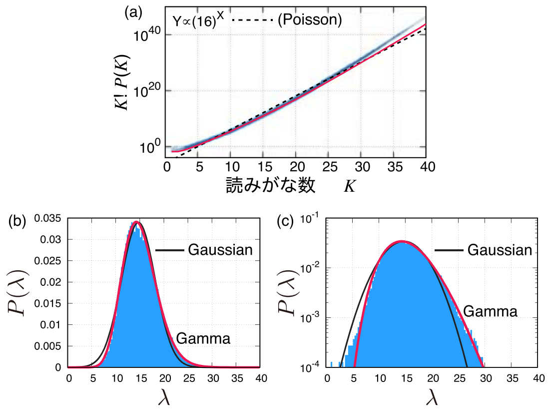

図 8 (a) 歌詞データにおける 1 行当たりの読みがな数の分布．(b)(c) 作品内での読みがな数のレート・パラメーターの分布と，ガンマ分布およびガウス分布によるフィット．

互いに独立に推定する必要がある．歪度と尖度を推定する標準的な方法は，3 次および 4 次モーメントに対する標本統計量を用いるものである．一方で，形質分布が共役分布（この場合はベータ分布）に従う場合は，形質分布の独立なパラメーターは 2 つだけのため，これらのパラメーターをフィットした後は，歪度と尖度（実際には全てのモーメント）は解析的に計算できる．ここでは具体的に，モーメント一致法を用いて，ベータ分布のパラメーターを標本平均と標本分散から決定するものとする．

ベータ分布のフィットを用いる方法の有効性を調べるため，ランダム抽出によって得られる部分データによりデータ量の減少を模倣して，この少量データから推定した歪度と尖度の値と元のデータから推定した値を比較する．推定誤差は，各データセットと各データ量に対して次の手順により求める：ランダムなデータ削減を 10 回繰り返し，それぞれに対して推定した統計量と元データから求めた標本統計量の誤差絶対値を平均する．図 6 に，標準的な方法により得られた平均推定誤差とベータ分布フィットを使う方法で得られたものをデータ量ごとにプロットしている．期待された通り，データ量が小さいほど後者の方法が相対的により効率的になっており，解析した音楽データの場合には，歪度と尖度の両方に対して，サンプル数が 70 以下の時に後者の方法がよりより効率的ということがわかった．形質分布が他の共役分布に近似的に従う場合にも同じ方法を適用できる．

5. 文芸と絵画データでの検証

3 節のモデルは，統計学習を通じて伝達される文化の一般的な性質に基づいて構成されている．これは，共役分布則が音楽特有の性質ではなく，その他の文化領域でも現れることを示唆している．ここでは，音楽以外の主要な芸術領域である文芸と絵画のデータを分析してこのことを検証する．

5.1 歌詞の読みがなの分析

文芸データの一例として，ここでは日本のポピュラー音楽の歌詞を分析する．インターネットで入手可能な公開データから，65,626 曲分の歌詞を抽出して分析に用いた．ポピュラー音楽の歌詞には英単語が含まれるものもあるが，この分析では英語文字は前処理で取り除いてから日本語部分のみを扱った．

まず，音楽データの音高の頻度に対応するものとして，歌詞の読みがなの頻度を調べた．MeCab により歌詞をひらがな表記の読みがなに変換し，各文字の各作品での頻度を求めてその分布を見た．図 7 に示す通り，多くの文字に対する分布が近似的にベータ分布に従うことがわかった．

次に，メロディーでの音符数に相当する量である読みがな数の分布を見る．歌詞は，通常 2 小節や 4 小節に対応する句ごとに改行して表示される．図 8(a) の様に，各行に含まれる読みがなの数 K の分布は，$5 \leq K \leq 30$ の領域でポアソン分布に近似的に従う．各作品内での読みがな数の分布をポアソン分布でフィットして得られる，レート・パラメーター λ の値は，その曲の行ごとの読みがなの全体的な数を表す．図 8(b)(c) の様に，この λ のデータセット内での分布を見ると，ガンマ分布に近似的に従うことがわかった．特に，ガウス分布ではフィットできない歪度があることや，ピーク値の約 2 桁下までガンマ分布でフィットできていることがわかる．

5.2 歌詞の単語ベクトル分布の分析

この節では，前節と同じ歌詞データを MeCab により単語分割したデータを用いて歌詞の内容を表す統計量を分析する．歌詞の分析には一般の文章の分析にも用いられる潜在ディリクレ配分法（LDA）が応用されているが [35,36]，ここではより高速で動く文章・単語埋め込みの方法を用いる [37]（論文の著者予測タスクでは，LDA よりも高性能であることも示されている）．この方法は word2vec と漸近的に等価である相互情報量の特異値分解 [38] を応用したものであり，文

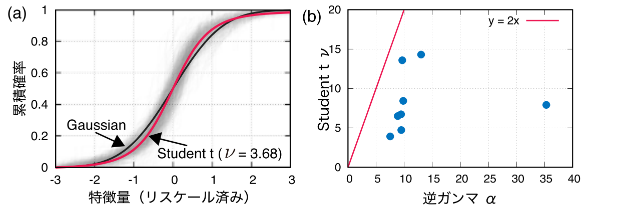

図 9 (a) 歌詞データにおける作品内の単語特徴量分布．(b) 作品内の平均と分散の分布から推定された t 分布の自由度 ν と逆ガンマ分布の α．ν が 100 より大きかった成分は表示していない．

章 d と単語 w の相互情報量 $Y_{dw} = \max\{P(w|d)/P(w), 0\}$ の特異値分解によりそれぞれを表現するベクトルが得られる．つまり，行列 Y をそれぞれランク K の行列 $D = (D_{dk})$ と $W = (W_{wk})$ により $Y \simeq DW^T$ と分解した際の D_{dk} が文章ベクトル，W_{wk} が単語ベクトルとなる．以下の分析では $K = 10$ とした．また，得られた文章ベクトルを調べると，空の単語列あるいはコンマやピリオドなどの非正則単語のみ含む歌詞からなるクラスターが見つかったため，このクラスターに属する歌詞を除いた 61,656 作品に対して分析を行った．

各作品の生産モデルとして，単語ベクトル W_{wk}（以下では単語特徴量と呼ぶ）の生成モデルを考える．各作品内の単語に対する単語特徴量は，各次元においておおよそ単峰のヒストグラムを持つが，サンプル数が少ないためその分布形をヒストグラムから正確に知ることは難しい．そこで，各次元のサンプルについて，平均 0 と分散 1 になるように標準化を行い，累積分布関数を求める．ランダムに抽出した 100 作品について，全ての特徴量次元で求めた累積分布を重ね合わせたものが図 9(a) である．まず，データ分布はピーク値を中心として左右対称であること，そして標準化を行なったあとの分布は全体的にかなり一致していることがわかる．また，図 9(a) に示す通りデータ分布はガウス分布に近いが，自由度 $\nu = 3.68$ とした Student の t 分布がよりよい近似を与えることがわかる．以上により，各作品の生産モデルは平均と分散の値をパラメーターとした t 分布で近似的に与えられることがわかった．t 分布はこれらのパラメーターに関して指数型分布ではないが，自由度が大きい場合はガウス分布に近いため，以下では平均と分散のデータセット内での分布とガウス分布に対する理論的結果を比較する．

3.4 節で示した様に，平均と分散をパラメーターとするガウス分布の共役分布はガウス・逆ガンマ分布で与えられる．

$$P(\mu, \Sigma) = \mathrm{Gauss}(\mu; \mu_0, \gamma\Sigma)\,\mathrm{InvGam}(\Sigma; \alpha, \beta). \quad (16)$$

但し，式 (15) では，$\mu_0 = \zeta_1$, $\gamma = (2Lu)^{-1}$, $\alpha = Lu + 3/2$, $\beta = Lu(\zeta_2 - \zeta_1^2)$ に対応する．この分布の特徴は，2 変数間の共分散が 0 であること，そして各変数に対する（もう一方を積分した時の）分布が以下の様に Student の t 分布と逆ガンマ分布になることである．

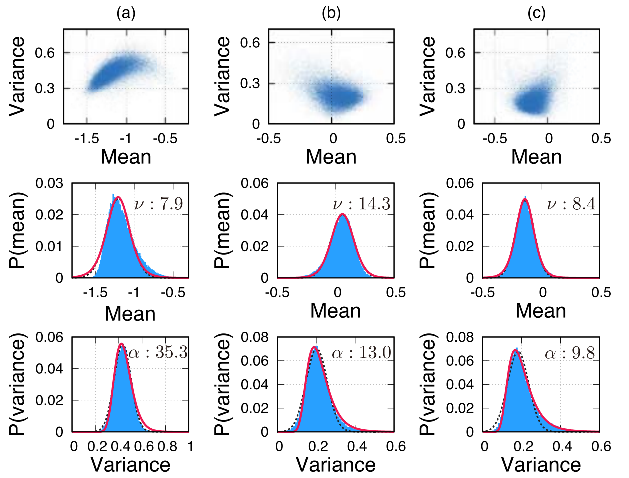

図 10 歌詞データにおける単語特徴量の作品内平均と分散の分布，および Student t 分布（平均）と逆ガンマ分布（分散）によるフィット．(a), (b), (c) はデータ全体での分散が第 1，5，9 番目に大きい特徴量成分に関する結果．点線はガウス分布のフィット曲線．

$$P(\mu) = \mathrm{St}(\mu; \mu_0, \gamma\beta/\alpha, 2\alpha), \quad (17)$$

$$P(\Sigma) = \mathrm{InvGam}(\Sigma; \alpha, \beta). \quad (18)$$

図 10 にデータから得られた作品内の特徴量の平均と分散の分布を示す．まず，2 次元の散布図を見ると，特に特徴量の第 1 成分 (a) では平均と分散の間の大きな共分散があることが確認される．(b) と (c) では分散の大きい方向にサンプルが広がって分布している様子も確かめられる．平均の分布は，(a) では大きな歪度を持っており，t 分布でのフィットも良くない．(b) と (c) は t 分布で良く近似されることがわかるが，推定された自由度 ν の値はいずれも大きく，ガウス分布でも比較的良いフィットになっている．分散の分布は，(a) では逆ガンマ分布よりガウス分布に近い形であり，(b) と (c) では右側の裾野が逆ガンマ分布により良くフィットできている．以上により，特徴量成分によって例外はあるものの，平均と分散の分布は共役分布のガウス・逆ガンマ分布に比較的合っていることがわかった．また，式 (17) と (18) では平均の t 分布の自由度 ν と分散の逆ガンマ分布の α に $\nu = 2\alpha$ の関係があることがわかる．データ分布から推定された ν と α の値を比較すると，図 9(b) の様に，$\nu \simeq \alpha$ に近い比例関係が見られ，理論から期待される関係とは異なる結果が得られた．

5.3 絵画データの色彩頻度統計量の分析

絵画データとして，インターネット上の公開サイト WikiArt [*1]から取得した約 13 万作品の絵画の画像ファイルからランダムに抽出した 10,000 ファイルを用いる．jpeg 形式の画像ファイルをリスケールおよびダウンサンプリングを行って，縦横各 50 ピクセルのビットマップデータに

[*1] https://www.wikiart.org/

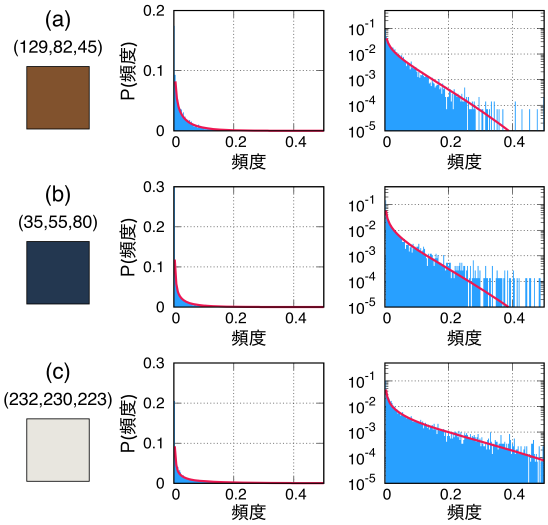

図 11 WikiArt データにおける代表色の作品内頻度とベータ分布によるフィット．

変換してから分析を行った．絵画の創作スタイルの重要な要素である色彩の特徴に注目する．

まず，各作品で使われている色の頻度統計量を分析する．絵画作品では描かれている物体ごとに色が違い，コントラストも多いため，RGB 色空間などの中での分布は複雑であることが多い．そこで，RGB 空間内で K 平均法によりクラスタリングを行い，30 個の代表色へ量子化した上で，各代表色の頻度を調べる．いくつかの代表色に対する作品内頻度の分布を図 11 に示す．頻度 0 にピークを最大値を持つヒストグラムが得られ，ベータ分布により 0 付近での急激な減衰と裾野における遅い減衰が精度良くフィットできることがわかった．

次に，色の組み合わせのスタイルを捉える統計量として，作品内での各代表色の頻度を大きい方から順に並べた時の分布を考える．図 12(a) のように，この順序統計量の外形は，頻度ランクが 3 から 20 の領域でおおまかに指数分布に従うことがわかる．作品内で使われる色の組み合わせは作品ごとに異なるが，この指数分布のレートパラメーター λ は使われる色の多様度を表す量として解釈できる（多数の色が面積均等に使われる作品は小さい λ を持ち，少数の色が大部分の面積を占める作品は大きな λ を持つ）．作品ごとに推定されたレートパラメーター λ の分布を図 12(b) に示す．図の通り，右側に裾が広がったこの分布は，共役分布であるガンマ分布に近似的に従うことがわかった．

5.4 考察

この節では文芸データと絵画データに対して，2 節で音楽データに対して行ったものと類似の統計的形質の分布の解析を行い，多くの場合に形質分布は共役分布として期待されるものに近似的に従うことが確認された．この結果は経験的な共役分布則が音楽データに限らず，様々な文化領域において見られる可能性を示している．また，理論モデルで表される統計学習を通した文化形質の伝達過程が多くの文化で共通した性質である可能性も示唆している．一方で，共役分布から外れた形質分布も観察されており，今後こうした分布形の違いがどういう状況で現れるかを調べることが重要だと考えられる．

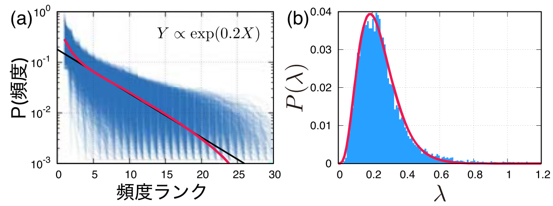

図 12 (a) WikiArt データにおける色頻度の順序統計量の分布．赤の太線は全作品での平均，黒の線は指数分布のフィット曲線を表す．(b) 各作品で推定した指数分布のスケールパラメーターの分布とガンマ分布によるフィット．

6. 議論

本稿では，統計学習を通して伝達される文化形質の分布が，斜行伝達があり選択圧がない場合に，形質の推定分散により決定される平衡分布に収束することを示した．この平衡分布は，データ生産モデルが指数型分布族の場合にその共役分布になり，定量的には，学習に用いられる文化生産物のサンプル数 L と斜行伝達の強さ u の積 Lu に依存する．導出された分布形は音楽・文芸・絵画データから抽出された経験的な分布により支持されることもわかった．

ここで調べたモデルでは，簡単のために 1 人の一次親と数多くの二次親の存在を仮定して，二次親の影響は “平均場” として扱った．一次親が複数いる場合にも，式 (1) は $\hat{\theta}_n^t$ の項をそれらの生産物の期待値の算術平均に置き換えることで拡張できる．もし一次親がランダムに選ばれるならば，そうしたモデルは連続形質に対する混合継承モデル [2] と類似したものになり，平衡における形質分布の分散は $O(D(\bar{\theta})/L)$ と見積もることができる．この場合には，u が小さい時には斜行伝達の効果は小さいと考えられる．これに対して，一次親として類似した形質を持つ個体を選ぶ傾向がある場合には，多数の一次親のモデルは 1 人の一次親のモデルに類似するため，斜行伝達の効果はより重要になる．一般には，文献 [9] の様な影響のネットワークを考えて，式 (1) の重み付け和をネットワーク上の分布に拡張することができる．そのようなモデルを用いて，ネットワーク構造により斜行伝達の影響の大きさがどう変化するかを調べることは興味深いであろう．

選択的に中立の場合，統計学習生成系の漸近的振る舞いは，ベイジアンエージェントの反復学習モデル [10,13] のそれと類似する．このモデルでは事前分布が平衡状態で維持される．よって，これらの大きく異なる伝達機構を区別するためには，選択圧下でのダイナミクスを調べる必要が

ある．モデル構築の観点からは，統計学習生成系は事前分布への依存を取り除き，選択圧がある場合にも柔軟な拡張ができる点で有用だと考えられる．

この研究の結果は，統計学習と文化進化の間の重要な関係性を示しており，これらを統計理論と興味深く結び付ける．式 (7) と (8) は，斜行伝達とベイズ学習の間にデータ生産モデルの幾何構造 [39] を反映した非自明な関係性があることを示している．式 (8) を満たす指数型分布族に対してはベイズ推定理論や極限定理を含む多くの統計理論的結果が知られており [32, 40]，それらを統計的性質を持つ文化形質の解析に適用することもできる．また，我々の結果は文化形質に対する適切なパラメトリックモデルを選択する際の理論的根拠を与えている点も特に重要である．例えば，広く用いられるガウス分布近似 [2, 3] は不適切な近似となり得る．

最後に，4 節でも議論したように，統計学習生成系は文化や知能の動的側面の研究において新しい実験分析の可能性を与える．ここで開発した分析法を音楽 [19]，言語 [18,41]，絵画 [20]，料理 [42] などの様々な文化領域に適用して，文化進化の普遍的性質を調べられる可能性がある．

謝辞　金子邦彦氏と竹内信人氏に有益な議論をしていただいたことに感謝する．本研究は，日本学術振興会科研費 19K20340, 21H03462, 21K12187，および 2021 年度京都大学リサーチ・ディベロップメントプログラム【いしずえ】から支援を受けた．

参考文献

[1] E. Szathmáry and J. Maynard Smith, Nature 374, 227 (1995).
[2] L. L. Cavalli-Sforza and M. W. Feldman, Cultural Transmission and Evolution (Princeton University Press, 1981).
[3] R. Boyd and P. J. Richerson, Culture and the Evolutionary Process (The University of Chicago Press, 1985).
[4] M. A. Nowak and D. C. Krakauer, Proc. Natl. Acad. Sci. 96, 8028 (1999).
[5] D. M. Abrams, H. A. Yaple, and R. J. Wiener, Phys. Rev. Lett. 107, 088701 (2011).
[6] J. R. Saffran, R. N. Aslin, and E. L. Newport, Science 274, 1926 (1996).
[7] J. D. Fernández and F. Vico, J. Artificial Intelligence Res. 48, 513 (2013).
[8] Y. LeCun, Y. Bengio, and G. Hinton, Nature 521, 436 (2015).
[9] G. J. Baxter, R. A. Blythe, W. Croft, and A. J. McKane, Phys. Rev. E 73, 046118 (2006).
[10] T. L. Griffiths and M. L. Kalish, Cog. Sci. 31, 441 (2007).
[11] F. Reali and T. L. Griffiths, Proc. R. Soc. B: Bio. Sci. 277, 429 (2010).
[12] C. Perreault, C. Moya, and R. Boyd, Evolution and Human Behavior 33, 449 (2012).
[13] B. Thompson, S. Kirby, and K. Smith, Proc. Natl. Acad. Sci. 113(16), 4530 (2016).
[14] B. Chazelle and C. Wang, J. Machine Learning Res. 20(29), 1 (2019).
[15] M. Marsili and Y.-C. Zhang, Phys. Rev. Lett. 80(12), 2741 (1998).
[16] C. Furusawa and K. Kaneko, Phys. Rev. Lett. 90(8), 088102 (2003).
[17] S. Meloni, J. Gońez-Gardenes, V. Latora, and Y. Moreno, Phys. Rev. Lett. 100(20), 208701 (2008).
[18] J.-B. Michel et al., Science 331, 176 (2011).
[19] M. Mauch, R. M. MacCallum, M. Levy, and A. M. Leroi, R. Soc. Open Sci. 2, 150081 (2015).
[20] A. Elgammal, B. Liu, D. Kim, M. Elhoseiny, and M. Mazzone, Proc. AAAI Conference on Artificial Intelligence 32(1), 2183 (2018).
[21] E. Nakamura and K. Kaneko, Sci. Rep. 9, 15993 (2019).
[22] C. M. Bishop, Pattern Recognition and Machine Learning (Springer, 2006).
[23] F. Simonetta, F. Carnovalini, N. Orio, and A. Rodà, Proc. Audio Mostly 2018 on Sound in Immersion and Emotion (2018).
[24] 椎葉京一, 久保昭二 (編集), 日本のうた 第 3–9 集 (野ばら社, 1998/1999/2001/2004/2014).
[25] B. L. Sturm, J. F. Santos, O. Ben-Tal, and I. Korshunova, Proc. Conference on Computer Simulation of Musical Creativity (2016).
[26] M. J. Salganik, P. S. Dodds, and D. J. Watts, Science 311, 854 (2006).
[27] R. A. Bentley, C. P. Lipo, H. A. Herzog, and M. W. Hahn, Evolution and Human Behavior 28, 151 (2007).
[28] W. J. Ewens, Mathematical Population Genetics 1: Theoretical Introduction (Springer, 2012).
[29] S. Wright, Genetics 16, 97 (1931).
[30] J. F. Crow and M. Kimura, An Introduction to Population Genetics Theory (Harper and Row, 1970).
[31] A. V. Fend, Ann. Math. Stat. 30(2), 381 (1959).
[32] C. N. Morris, Ann. Stat. 10(1), 65 (1982).
[33] P. E. Savage, S. Brown, E. Sakai, and T. E. Currie, Proc. Natl. Acad. Sci. 112, 8987 (2015).
[34] E. Nakamura, preprint, arXiv:2102.01465.
[35] S. Sasaki, K. Yoshii, T. Nakano, M. Goto, and S. Morishima, Proc. ISMIR, 585 (2014).
[36] 佃洸摂, 石田啓介, 後藤真孝, 情報処理学会インタラクション 2018, INT18012, 96 (2018).
[37] 持橋大地, 言語処理学会第 27 回年次大会, B2-2 (2021).
[38] O. Levy and Y. Goldberg, Adv. Neural Inf. Proces. Sys. 27, 2177 (2014).
[39] S. Amari, Information Geometry and Its Applications (Springer, 2016).
[40] C. N. Morris, Ann. Stat. 11(2), 515 (1983).
[41] T. L. Griffiths and M. Steyvers, Proc. Natl. Acad. Sci. 101(suppl. 1), 5228 (2004).
[42] M. Caraher, P. Dixon, T. Lang, and R. Carr-Hill, British Food Journal 101(8), 590 (1999).