

情報カスケードと一般化ポリア壺過程

Information cascade and generalized Pólya urn

守 真太郎
Mori Shintaro

北里大学理学部
School of Science, Kitasato University
shintaro.mori@gmail.com, <http://202.24.143.74/mori/>

久門 正人
Hisakado Masato

金融庁
Financial Services Agency
masato.hisakado@fsa.go.jp

日野 正文
Hino Masafumi

日本電気株式会社
NEC Corporation
masafumi.hino@gmail.com

高橋 泰城
Takahashi Taiki

北海道大学文学研究科
Department of Behavioral Science, Hokkaido University
taikitakahashi@gmail.com

キーワード: information cascade, phase transition, nonlinear Pólya urn, experiment

概要

情報カスケードの集団実験で起きるマクロな変化は非線形ポリア壺過程の写像 $q(z)$ の安定固定点の個数の変化として理解できる。このマクロな変化が相転移かどうか、また、相転移の場合、秩序変数、普遍クラスについて議論する。結論は、(1) 相転移であり、秩序変数は最初の被験者と後の被験者の選択の相関関数 $C(t)$ の極限值 $c = \lim_{t \rightarrow \infty} C(t)$ である。 c は最初の被験者の選択の後の被験者全体への影響（ドミノ効果）がパーコレーションするかどうかを記述する。(2) $C(t)$ は吸収状態相転移の秩序変数と類似のスケーリング構造を持ち、臨界指数間にスケーリング関係式が成立する。(3) 相転移は一般に不連続転移だが、系が Z_2 対称性を持つ場合には連続転移となる。実験でのマクロな変化は系が非対称な場合に対応し、不連続転移である。(4) 連続転移の普遍クラスは秩序変数の臨界指数 β は $\beta = 1/2$ でイジング平均場モデルと同じである。情報カスケードでの非平衡相転移は臨界点での相関関数 $C(t)$ が $C(t) \propto \log t^{-1/2}$ と振る舞うことで特徴づけられる。ただし、(3)(4) は臨界点での $q(z)$ が touch point で対角線と接する場合である。

1. はじめに

情報カスケードとは多数のヒトが順番に過去のヒトの選択情報を参照しながら選択するとき、各自の持つ自己情報の示す選択肢ではなく、多数派の選択肢を選ぶ傾向のことをいう [BHW 92]。自己情報が不確実な場合、多数派の選択肢が正解であることを予想して選択していると考えて合理的ハーディングとも呼ばれる。情報カスケードの起きる状況とは、ある選択肢の選択者数が他の選択肢を大きく上回り、一人の選択では多数派が覆らない状況である。そのため、一旦そうした状態になると多数派の選択肢が選ばれる連鎖が起きる [Anderson 97]。

こうした連鎖が無限に続くのか、それともいつれ終わるのかについては議論がある [Goeree 07]。特に、ある時点で間違った選択肢が多数派になりその連鎖が続いても、十分多数のヒトが選択すると正しい選択肢が多数派になるのかどうかである。これを自己修正性という。現実の情報カスケード、例えばバブル、流行、デマなどでは外部からの情報により多数派の選択の連鎖が壊れる

こと、壺の選択を用いた実験室での集団実験でも多数派の選択肢の連鎖が終わることなどから連鎖の状態は脆弱であるといわれてきた [BHW 92]。一方、新商品や新技術の採用でのある新商品や新技術が市場を占有するカスケード現象は Lock-In と呼ばれ安定であると考えられている [Arthur 90]。我々は二択のクイズを用いた情報カスケード実験を行い、クイズの難易度の変化により系のマクロな振る舞いに変化が起き、相転移であることを報告した [Mori 12, Mori 13]。クイズが簡単な場合、多数派が間違っているカスケードは脆弱であり、十分多数の人が選択すれば正しい選択肢が多数派になる。一方、クイズが難しい場合、そうした修正は必ず起きるわけではなく、ある時点での間違った選択肢が多数派となったとき、その状態を永遠にキープする確率は正である。つまり、多数派が間違っているカスケード状態は脆弱ではなく安定である。

こうしたマクロな変化は非線形ポリア壺過程での写像 $q(z)$ の変化として理解できる [Hisakado 11, Hisakado 12]。非線形ポリア壺過程とは、写像 $q(z) : [0, 1] \rightarrow [0, 1]$

と次の確率法則に従う二値の確率過程 $X(t) \in \{0, 1\}, t = 1, 2, \dots$ のことである [Hill 80].

$$\Pr(X(t+1) = 1 | z(t) = z) = q(z) \quad (1)$$

ここで、 $z(t)$ は $X(1), \dots, X(t)$ の 1 の比率である.

$$z(t) = \frac{1}{t} \sum_{s=1}^t X(s)$$

非線形ポリア過程は $X(t+1)$ が $z(t)$ を通して過去の影響を受ける. $q(z)$ の固定点 $z_* \in \{z | q(z) = z\}$ のうち、 z_* の近傍で $q(z) - z$ が正から負になるものを安定固定点、負から正になるものを不安定固定点と呼ぶ (図 1 参照). $z(t)$ が安定固定点に収束する確率は正、不安定固定点に収束する確率はゼロである [Hill 80]. 安定固定点があつしかない場合、 $z(t)$ はその唯一の固定点に収束する. 一方、2 つ以上ある場合、どの固定点に収束する確率も正である. 情報カスケードの集団実験でも $q(z)$ の振る舞いを調べるにより安定固定点の個数が変化することが確かめられた [Mori 12, Mori 13].

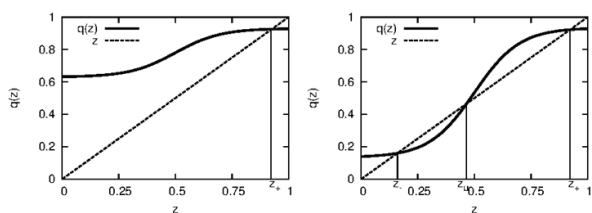


図 1 $q(z)$ のプロット. 左は安定固定点が $z = z_+$ にひとつ. 右は安定固定点が $z = z_+$ に 2 個と不安定固定点が $z = z_-$ にひとつ.

では、非線形ポリア壺での安定固定点の個数の変化は相転移なのだろうか? 統計物理学では相転移はマクロな系のマクロな物理量の示すパラメータ依存性の非解析性として定義される. パラメータ変化によりある相から別の相に変化するときマクロな物理量に不連続性や微分不連続性などの非解析性が現れる. 前者を不連続転移、後者を連続転移と呼ぶ. 秩序変数と呼ばれる物理量が各相を特徴付け、連続転移においては相関距離と呼ばれる長さのスケールが現れ、マクロな物理量の特異性が相関距離でスケールされ、パラメータ依存性が相関距離の同時関数となる. また、同時関数の構造、特に臨界指数が相転移を分類する普遍クラスを決める. ここでは、非線形ポリア壺のパラメータ変化によるマクロな変化の秩序変数とその構造を解析し普遍クラスを明らかにする [Mori 15, Mori 15a]. また、それに基づく実験データの解析手法について紹介する [Mori 15b, Hino 15].

2. 非線形ポリア壺と安定状態

非線形ポリア壺の安定状態 z を $z(t)$ が z に収束する確率が正と定義する. 写像 $q(z)$ の安定固定点以外にも

touch point と呼ばれる z 軸上の孤立点 z_t が安定状態になることもある [Pemantle 91]. ここで、touch point とは、 z_t で $q(z)$ が対角線に交叉しない不動点 $q(z_t) = z_t$ で、 z_t の近傍で対角線より上、または下にある孤立点である (図 2 参照). $q(z)$ が z_t の左側で上にある場合、 z_t は安定状態となることがある. z_t より右では $q(z) > z_t$ となり、 $z(t)$ は平均して右側に動く. z_t より左にある場合でも $q(z) > z_t$ より、 $z(t)$ は平均的に右に動くが、問題は z_t を確率 100% で通過するか、それとも通過しない確率が有限かどうかである. 直感的に言えば、 $q(z)$ と z の差が大きければ $z(t)$ を右に押す力が強く z_t を通過する. 逆に小さいなら z_t を通過する時間が無限大となり、 z_t は安定状態になる. z_t が安定かどうかは $q(z)$ の z_t より左側の振る舞いで決まる. その条件は $q(z)$ の z_t における左微分で決まり、それが $1/2$ 未満なら不安定、 $1/2$ より上なら安定である.

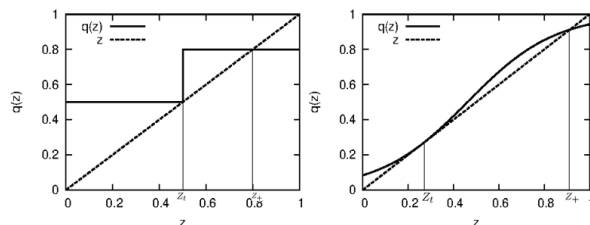


図 2 touch point があるときの $q(z)$ のプロット. 左のデジタル型の $q(z)$ では touch point z_t は不安定. 右のイジング型の $q(z)$ では touch point z_t は安定.

図 2 は $q(z)$ が閾値 $z = 1/2$ を境にヘビサイド関数的に変化する場合 (左図) となめらかに単調増加する場合 (右図、後のイジング型に対応) の touch point の様子を示している. 前者の場合、 $q(z)$ の z_t での左微分が 0 となり、 z_t は不安定である. 後者の場合、 $q(z)$ は対角線に接し左微分は 1 なので、 z_t は安定である.

2.1 秩序変数 c

ポリア壺過程の安定状態は安定固定点と前節で述べた安定な touch point である. では、 $q(z)$ のパラメータ変化で安定状態の個数が変化するとき何が起きるのだろうか? 安定状態が z_+ に 1 個しかない場合、 $z(t)$ は初期条件 $z(1) = X(1)$ に依存せず z_+ に収束する. 一方、安定状態が $z_+ > z_-$ の 2 個の場合、 $z(t)$ がそれぞれに収束する確率は初期条件に依存するであろう. $z(1) = X(1) = 1$ なら z_+ に収束する確率が大きく、 $z(1) = X(1) = 0$ なら z_- に収束する確率が大きいと考えられる. $z(t)$ が z_{\pm} に収束する確率を $X(1)$ 依存性を考慮して $p_{\pm}(x)$ と書くことにする.

$$p_{\pm}(x) \equiv \Pr(z(t) \rightarrow z_{\pm} | X(1) = x).$$

初期条件 $X(1)$ の違いによる $p_+(x)$ の差を考えてみる.

$$p_+(1) - p_+(0)$$

この確率の差は $X(1)$ の値の違いでの z_+ に収束する確率の差であり, $X(1)$ と $X(t+1)$ の相関関数 $C(t)$ の極限值 c に比例している.

$$c = (q(z_+) - q(z_-)) \cdot (p_+(1) - p_+(0)) \quad (2)$$

$$c \equiv \lim_{t \rightarrow \infty} C(t)$$

$$C(t) \equiv \text{Cov}(X(1), X(t+1)) / \text{Var}(X(1)) \quad (3)$$

ここで, A の期待値, 分散を $E(A)$, $\text{Var}(A)$, A, B の共分散 $\text{Cov}(A, B)$ を $\text{Cov}(A, B) = E(AB) - E(A) \cdot E(B)$ と定義している. 一般に $p_+(1) > p_+(0)$ が成立するので, $c > 0$ である. 一方, 安定状態がひとつしかない場合, $X(1) = x$ の値によらず $p_+(x) = 1$ となるため $c = 0$ となる.

こうして, 安定状態の個数が 1 から 2 に変化するとき, ゼロ ($c = 0$) から正の値 ($c > 0$) に変化するマクロな物理量 c が定義された. では, c は相転移の秩序変数と考えてよいのだろうか? c の意味を明らかにするために, 非平衡相転移の一種である吸収状態相転移の秩序変数の振る舞いと類似を説明する. 吸収状態相転移とは, 確率過程において吸収状態と呼ばれる, その状態になると変化しなくなる状態に確率 100% で変化するか, 吸収状態に変化せず, 他の状態にとどまる確率が正のままかの 2 つのマクロな状態間の非平衡相転移である [Hinrichsen 00]. この相転移では, 時刻 t までに吸収状態に移らない確率 $\rho(t)$ の極限值が秩序変数となる. $\rho(t)$ は時間軸上の相関時間 ξ と普遍関数と呼ばれる $x = 0$ で正則な関数 $g(x)$ を用いて

$$\rho(t) = t^{-\alpha} g(t/\xi) \quad (4)$$

とスケールすると考えられている. これを現象論的スケールリング則と呼ぶ. そして, $\rho(t)$ の極限值が有限であることから, $g(x)$ の $x \rightarrow \infty$ での漸近的な振る舞いは

$$g(x) \sim x^\alpha$$

となることが分かる. すると,

$$\rho = \lim_{t \rightarrow \infty} \rho(t) = \xi^{-\alpha}$$

という関係が成立することが分かる. 系があるパラメータ p に対して $p = p_c$ で連続転移するとき, $\Delta p = |p - p_c|$ を用いて $\rho \propto \Delta p^\beta$, $\xi \propto \Delta p^{-\nu_{||}}$ で $\nu_{||}, \beta$ を定義すると

$$\beta = \alpha \cdot \nu_{||}$$

が成立する. これを現象論的スケールリング関係と呼ぶ.

非線形ポリア壺の場合 $z(t) = z_+$ となっても状態変化が止まるわけではないので z_+ は吸収状態ではない. そのため非線形ポリア壺のマクロの安定状態の個数の変化

は厳密には吸収状態相転移ではない. しかし, $c = 0$ の場合 $t \rightarrow \infty$ の極限で $z(t)$ は初期状態によらず確率 100% で z_+ に収束し, $c > 0$ の場合 $z(t)$ が z_+ に収束する確率は 100% 未満である点は c が秩序変数であることを示唆する. そこで, $C(t)$ のスケール関係を調べてみる [Mori 15, Mori 15a].

2・2 厳密に解ける場合: デジタル型

厳密に解ける非線形ポリア壺過程の $C(t)$ の構造を見つめることにする. 写像 $q(z)$ が 2 個のパラメータ p, q_* を用いて次の関数で与えられる場合である.

$$q(z) = (1-p) \cdot q_* + p \cdot \theta(z - \frac{1}{2})$$

$X(t+1)$ は $z(t)$ が $1/2$ より大か小という多数決ルールで過去の影響を受けるモデルである. $z(t) > 1/2$ なら $X(t+1)$ が 1 となる確率は $(1-p) \cdot q_* + p$, $z(t) < 1/2$ なら確率は $(1-p) \cdot q_*$ である. デジタル的な影響のモデル化なのでデジタルモデルと呼ぶ. $p \in [0, 1]$ は過去の影響の強さを表すパラメータである. また, q_* は $p = 0$ の場合, つまり過去の影響がない場合に $X(t+1) = 1$ となる確率を表す. 正しい選択肢を 1, 間違った選択肢を 0 で表すとすると, q_* は独立に選択したときの正答率である. 以下, $q_* \geq 1/2$ と仮定する.

このモデルの安定状態は $p \leq p_c(q) = 1 - 1/2q_*$ のとき, $z_+ = (1-p)q_* + p$ に 1 個, $p > p_c(q_*)$ のとき, z_+ と $z_- = (1-p)q_*$ に 2 個存在する. $p = p_c(q_*)$ のとき, $z = 1/2$ に touch point が存在するが, $q(z)$ の左微分は 0 なので不安定である. そのため, $p = p_c(q_*)$ でも $z(t)$ は z_+ に収束する (図 2 の左). $q_* > 1/2$ のとき $p_c > 0$ となり, $C(t)$ の極限值 $c(q_*, p)$ は, $p \leq p_c(q_*)$ で 0, $p > p_c(q_*)$ で正の $p = p_c(q_*)$ で微分不連続な関数となり, 連続転移を示唆する. $q_* = 1/2$ のとき $p_c = 0$ となり, 系は $p = 0$ で $z_+ = 1/2$, $p > 0$ で $z_{\pm} = 1 \pm p/2$ に 2 個の安定状態を持つ.

$q_* > 1/2$ のとき $C(t)$ は, 相関距離 $\xi(q_*, p)$, 普遍関数 $g(x)$ を用いて

$$C(t) = b(q_*) t^{-1/2} g(t/\xi(q_*, p)) \quad (5)$$

と書くことができる. ここで, $b(q_*), \xi(q_*, p)$ は

$$b(q_*) = \sqrt{\frac{8}{\pi}} \left(\frac{2q_* - 1}{4q_* - 1} \right),$$

$$\frac{1}{\xi(q_*, p)} = -\log(4(p + (1-p)(1-q_*))((1-p)q_*))^{\frac{1}{2}}$$

で定義される. 普遍関数 $g(x)$ は, $p > p_c(q_*)$, $p < p_c(q_*)$ で 2 つの分岐を持ち, 次のように書ける.

$$g(x) = \begin{cases} \sqrt{4\pi x} + \frac{x^{1/2}}{2} \int_x^\infty u^{-3/2} e^{-u} du, & p > p_c(q_*) \\ \frac{x^{1/2}}{2} \int_x^\infty u^{-3/2} e^{-u} du, & p < p_c(q_*) \end{cases} \quad (6)$$

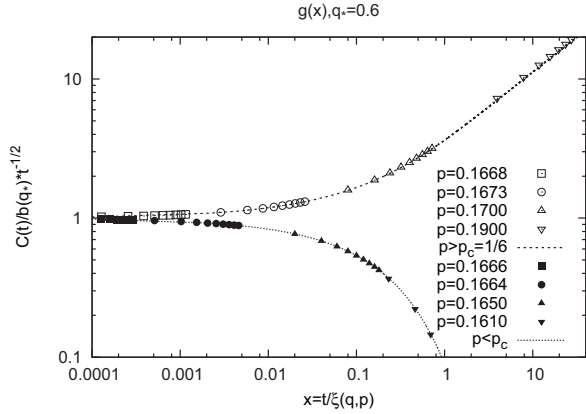


図3 $C(t)/b(q_*)t^{-1/2}$ を $t/\xi(q_*,p)$ に対してプロット. $q_* = 0.6$ とし $t \leq 10^5$ までマスター方程式を数値的に積分して求めた $C(t)$ を用いている. $p > p_c(q_*)$ を \square, \circ などの記号で, $p < p_c(q_*)$ を \blacksquare, \bullet などの記号で示している. $q_* = 0.6$ とし $10^4 \leq t \leq 10^5$ の t に対して計算している. 破線, 点線は式6を表している.

$q_* = 0.6$ の場合に $t = 10^5$ までマスター方程式を数値積分して $C(t)$ を評価し, 式(5)を用いて $g(x)$ をプロットしたものが図3である.

$C(t)$ は吸収状態相転移での秩序変数 $\rho(t)$ と同じスケールリング構造を持つことが分かる. つまり, デジタルモデルの $q_* > 1/2$ での連続転移の構造は吸収状態相転移での秩序変数と同じであり, $C(t)$ の極限值 c が秩序変数であることを意味する.

2.3 イジング型の場合

touch point で $q(z)$ が対角線に接し安定状態となる場合を考える. この場合, $q(z)$ の微分は1となり安定状態である(図2の右). $q(z)$ の選択には任意性があるが, 実験での応答関数から2個のパラメータ J, h を用いて次の関数で与えられるイジング型のものを用いる.

$$q(z) = \frac{1}{2}(\tanh(J(2 \cdot z - 1) + h) + 1).$$

ここで h は系の非対称性のパラメータである. $h = 0$ の場合, 系は $q(z) = 1 - q(1 - z)$ という Z_2 対称性を持つ. $J = 0$ の場合, $X(t+1) = 1$ となる確率は $(\tanh h + 1)/2$ となり, $h = 0$ で $1/2$ の h の増加関数となる. また, J は過去の選択情報 z の $X(t+1)$ への影響を決めるパラメータである. 以下, $h > 0, J > 0$ とする.

写像 $q(z)$ の固定点 z_* の満たす条件 $z_* = q(z_*)$ は, $S(t) = 2 \cdot X(t) - 1$ という変数変換を行うことで $m = 2z_* - 1$ に対する方程式となり,

$$m = \tanh(J \cdot m + h)$$

となる. これは平均場イジングモデルの磁化 m の満たす自己無撞着方程式と同じものである. この方程式は $J_c(h)$ をしきい値として, $J < J_c(h)$ なら $m = m_+(h)$ に1個の解,

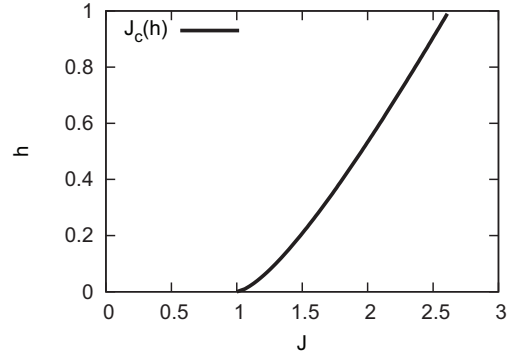


図4 (J, h) 面での相図. $J < J_c(h)$ では安定状態はひとつ. $J > J_c(h)$ では安定状態は2個. $h = 0, J = J_c(0) = 1$ では, $z = 1/2$ に安定状態の touch point. $h > 0, J = J_c(h)$ では安定固定点 $z_+ > 1/2$ と安定な touch point z_t の2個の安定状態.

$J > J_c(h)$ は $m = m_{\pm}, m_u$ に3個の解を持つ. 特に $h = 0$ の場合, $J_c(0) = 1$ で, $J < J_c(0)$ なら $m_+ = 0, J > J_c(0)$ なら $\pm|m_+|$ に2個の安定固定点と $m = 0$ に不安定固定点を持つ. z 座標で考えると, $J \leq 1$ のとき $z_+ = 1/2$ が安定固定点. $J > 1$ のとき, $z_+, z_- = 1 - z_+$ に2個の安定固定点と $z = 1/2$ の不安定固定点を持つ. $J = 1$ で z_+ は touch point であるが, 安定状態はこれのみなので, $c = 0$ となる. $J > 1$ で J が1より大きくなると, z_{\pm} は $1/2$ から連続的に離れるので, $p_+(1) - p_+(0)$ が J について連続的に振る舞うなら c も J の連続関数となり, $J < 1$ から $J > 1$ の変化は連続転移と考えられる. 実際, 以下で示す $C(t)$ のスケールリング則から連続転移であることが分かる. 一方, $h > 0$ のとき, $J < J_c(h)$ では安定固定点は z_+ のみ, $J > J_c(h)$ は z_{\pm} に2個の安定固定点と1個の不安定固定点を持つ. $J = J_c(h)$ では, $z = q(z)$ は z_+ の解の他に z_t に touch point を持ち, ともに安定である. z_+ に収束する確率 $p_+(x)$ は初期状態 $X(1) = x$ に依存すると考えられるので $c > 0$ となり, $J < J_c(h)$ から $J > J_c(h)$ の変化は不連続転移となる.

$h = 0$ の場合の $C(t)$ のスケールリング則を調べる. ただし, 通常の臨界現象のスケールリングとは根本的に異なる. なぜなら, $J < J_c$ での $C(t)$ と $J > J_c$ での $C(t) - c$ の t 依存性は t のべき則に従うからである. デジタルモデルの場合, $C(t) - c$ は指数関数的に減衰した. 一方, イジング型のモデルの $C(t)$ の漸近形は, 極限值 c と係数 c' を用いて

$$C(t) \simeq c + c' \cdot t^{l-1} \quad (7)$$

と書くことができる. l を $\{q'(z_+), q'(z_-)\}$ の大きい方の値と定義する.

$$l = \text{Max}(l_+ \equiv q'(z_+), l_- \equiv q'(z_-)).$$

$h \neq 0$ の場合, l は $J = J_c(h)$ で不連続に変化する. $J < J_c(h)$ では $l = l_+ < 1, J = J_c(h)$ で touch point z_- が現れると $l_- = 1$ より $l = 1$ となるためである. $h = 0$ の場

合, $J \leq J_c(0) = 1$ で $l = J$, $J \geq J_c(0)$ で $1 - l \propto J - 1$ と振る舞い, l は J の連続関数となる.

このように, $C(t)$ は秩序変数 c を除くと基本的に t のべき関数であるため, 吸収状態相転移での現象論的スケールリング関係 (4) は成立しない. この場合成立するのは, $\xi = 1/(1-l)$ を $\log t$ スケールでの相関距離と見做した次のスケールリング関係である.

$$C(t) = \log t^{-\alpha'} \cdot g((1-l) \log t) \quad (8)$$

このとき, ξ は $\Delta J = |J - J_c(0)|$ に対し, $\xi \propto \Delta J^{-1}$ と振る舞うので, $\nu_{||} = 1$ である. あとは現象論的スケールリングの議論を用いて, $\alpha', \nu_{||}$, および β の間に

$$\beta = \alpha' \cdot \nu_{||}$$

が成立することが分かる. これがイジング型, より一般には touch point で写像 $q(z)$ が対角線に接し, 系が Z_2 対称性を持つ場合の連続転移において成立するスケールリング関係式である.

$\nu_{||}$ 以外の指数 α', β を求めるために数値計算を行う. イジング型のモデルでのマスター方程式を数値積分し $C(t)$ を評価して $\alpha' = 1/2$ と仮定し, (8) を用いて評価した $g(x)$ をプロットしたものが図 5 である.

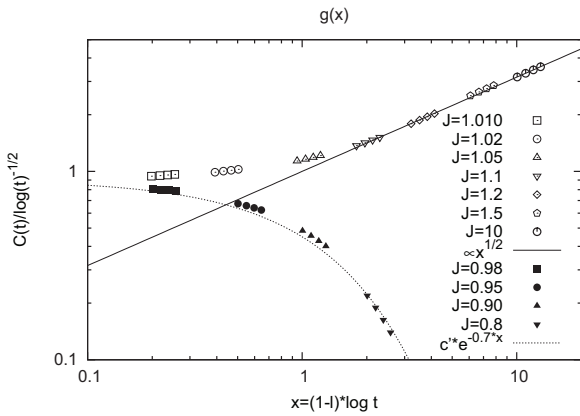


図 5 $C(t)/\log t^{-1/2}$ を $(1-l) \cdot \log t$ に対してプロット. $t \leq 10^6$ までマスター方程式を数値的に積分して求めた $C(t)$ を用いている. $J > J_c(0)$ を \square, \circ などの記号で, $J < J_c(0)$ を \blacksquare, \bullet などの記号で示している.

$J < J_c(0) = 1$ のときと $J > J_c(0)$ のときの $g(x)$ は $J = J_c(0)$ で $g(0)$ となり, そこから別の分岐を描く. $J < J_c(0)$ の場合, $g(x)$ は x とともに指数関数的に減衰する. 一方, $J > J_c(0)$ のとき, x が十分大きくなると $x^{1/2}$ に従って増加していることが分かる. これは, $J > J_c(0)$ では $c > 0$ より, $c \cdot \log t^{1/2}$ と増加することから来る. 異なる t, J のデータ $C(t)$ が $(1-l) \cdot \log t$ と $C(t)/\log t^{-1/2}$ により同じカーブに乗る. よって (8) のスケールリングが $\alpha' = 1/2$ で成立している. スケールリング関係式 $\beta = \alpha' \cdot \nu_{||}$ より, $\beta = 1/2$ となる. この指数は平均場イジングの指数と一

致している. ただし, このことは平均場イジングの平衡相転移とポリア壺の相転移が同じ物理という意味では全くない. 平衡相転移では, $J < J_c(0)$ に対応する無秩序相では初期状態の記憶を完全に失った状態が平衡状態であるが, ポリア壺過程ではべき関数で初期状態の記憶を永遠に保持している. β が一致したのはイジング型の写像 $q(z)$ を選んだことによるものである.

3. 情報カスケード実験のデータ解析手法

ここでは, 非線形ポリア壺の相関関数 $C(t)$ の漸近形 (7) に基づいた情報カスケード実験のデータ解析手法について述べる [Mori 15b, Hino 15]. $J = J_c(h)$ の場合, $C(t)$ は

$$C(t) \simeq c + c' \log t^{-\alpha'}$$

と $\log t$ のべき関数として振る舞うが, 集団実験で $J = J_c(h)$ にパラメータを設定することは難しく考えない [Mori 15a]. 相転移を検出するには c を評価する必要がある. なぜなら c は秩序変数であり, $c > 0$ と $c = 0$ が各相を特徴づけるからである.

実験データは, コントロールされた条件下での多数の二値の時系列データである. サンプル数を I , 時系列の長さを T とし, データを $\{X(i, t), X(i, t) \in \{0, 1\}, i = 1, \dots, I, t = 1, \dots, T\}$ で記述する. 問題はデータを用いて評価した自己相関関数 $C(t)$ の極限值 c およびべき指数 l の推定方法である. 実験データではサンプル数 I , サンプル長 T とともに大きくはなく, 後に解析する実験データではサンプル数 I は 40 から 200, T は 60 から 70 程度である.

まず, $C(t)$ の評価は共分散の式 (3) ではなく, 次の条件付き確率の差の表式を用いる.

$$C(t) = \Pr(X(t+1) = 1 | X(1) = 1) - \Pr(X(t+1) = 1 | X(1) = 0). \quad (9)$$

実験 II はサンプル数も多いので, 条件付き確率 $\Pr(X(t+1) = 1 | X(1) = x)$ の推定値 $q_x(t+1)$ とその標準誤差 $\Delta q_x(t+1)$ は

$$q_x(t+1) = \frac{\sum_i \delta_{X(i,1),x} X(i, t+1)}{N_x},$$

$$N_x = \sum_i \delta_{X(i,1),x},$$

$$\Delta q_x(t+1) = \sqrt{\frac{q_x(t+1)(1 - q_x(t+1))}{N_x}}, \quad (10)$$

とする. 実験 I は特に 70%, 80% のケースでサンプル数が少ないため, q_x の評価にベイズ統計の事後分布の期待値を用いている. $q_x(t+1)$ の事前分布として一様分布 $\beta(1, 1)$ を仮定する. $X(i, 1) = x$ の各条件 (サンプル数はそれぞれ N_x) での $X(i, t+1)$ の 1, 0 の個数を $N_1(t+1), N_0(t+1)$ のとき, 事後分布は β 分布 $\beta(N_1(t+1) +$

1, $N_0(t+1)+1$) に従う。 $q_x(t+1)$ の推定値および標準誤差は事後分布の期待値および標準偏差とする。

$$q_x(t+1) = \frac{N_x(t+1)+1}{N_x+2}$$

$$\Delta q_x(t+1) = \sqrt{\frac{q_x(t+1)(1-q_x(t+1))}{N_x+3}}$$

$C(t)$ およびその標準誤差 $\Delta C(t)$ は $X(1) = 1$ および $X = 0$ のサンプルの独立性から

$$C(t) = q_1(t+1) - q_0(t+1),$$

$$\Delta C(t) = \sqrt{\Delta q_1(t+1)^2 + \Delta q_0(t+1)^2}, \quad (11)$$

で評価する。

サンプル数 I が十分大きく標準誤差 $\Delta C(t)$ が小さいなら、 $C(t)$ の漸近形 (7) を仮定してフィットすることにより c を評価することは可能かもしれない。実際、マスター方程式を数値的に積分し、十分大きな t まで厳密に $C(t)$ を計算することは容易である。しかし、実験データのサンプル数 I は限られているため $\Delta C(t)$ は大きく、フィッティングによる秩序変数 c の評価は難しい。そこで、 c の推定量として $C(t)$ の積分量を用いることを考える。それは積分相関時間 τ を t で割った τ_t である。また、 $c=0$ の場合の $C(t)$ の主要項は $c' \cdot t^{l-1}$ だが、そのべき指数 l を評価する推定量として τ_t に加えて 2 次モーメント相関時間 ξ を t で割った ξ_t を用いる。 l は $C(t)$ の主要項のべき指数なので、 $c > 0$ の場合 $l = 1$ と考える。 τ_t, ξ_t は、 $C(t)$ のモーメント $m_n(t)$ を用いて次のように定義される。

$$\tau_t = \frac{1}{t} m_0(t)$$

$$\xi_t = \sqrt{m_2(t)/m_0(t)}$$

$$m_n(t) = \sum_{s=0}^{t-1} C(s)(s/t)^n \quad (12)$$

$C(t)$ の漸近形を用いて τ_t の漸近形を求めると、

$$\tau_t(t) \simeq c + \frac{c'}{t} t^{l-1}. \quad (13)$$

$\xi_t(t)$ の漸近形は

$$\lim_{t \rightarrow \infty} \xi_t(t) = \begin{cases} \sqrt{\frac{l}{l+2}}, & c = 0 \\ \sqrt{\frac{1}{3}}, & c > 0. \end{cases} \quad (14)$$

となる。

これら推定量の標準誤差は、 $\Delta C(s), s < t$ が互いに独立であるとの仮定のもとで次の式で評価する。

$$\Delta \tau_t = \frac{1}{t} \Delta m_0(t),$$

$$\Delta \xi_t = \sqrt{\xi_t \left(\frac{\Delta m_2(t)}{2m_2(t)} + \frac{\Delta m_0(t)}{2m_0(t)} \right)}$$

$$\Delta m_n(t) = \sqrt{\sum_{s=1}^{t-1} \Delta C(s)^2 (s/t)^{2n}}. \quad (15)$$

表 1 実験 I(EXP-I), 実験 II(EXP-II) の概略。実験 I は 2012 年および 2014 年に北大文学部, 北里大理学部で実施。実験 II は 2013 年, 2014 年に北里大理学部で実施。実験 I では、ノーヒントでの正答率 $Z_0(T)$ で問題を分類。 $45\% \leq Z_0(T) < 55\%$, $55\% \leq Z_0(T) < 65\%$, $65\% \leq Z_0(T) < 75\%$, $75\% \leq Z_0(T) < 85\%$ 。実験 II では、被験者が壺 X からランダムに取り出した玉の色と壺 X の種類が一致する事後確率 q を $q \in \{5/9, 6/9, 7/9\}$ で実施。

EXP-I				
$z_0(T)$	I	N_1	N_0	T_{avg}
50	73	35	38	71.2
60	79	55	24	67.7
70	46	37	9	70.7
80	43	38	5	63.2
EXP-II				
q	I	N_1	N_0	T_{avg}
5/9	200	115	85	63
6/9	200	99	101	63
7/9	200	156	44	53.7

$\Delta C(s)$ が互いに独立として $\Delta m_n(t)$ の評価しているが、 $\Delta m_2(t)$ と $\Delta m_0(t)$ は完全にシンクロするとして $\Delta \xi_t$ を大きめに評価している。

3.1 秩序変数 c とべき指数 l の評価

以下、2種類の情報カスケード集団実験のデータを用いて秩序変数 c とべき指数 l を評価する。実験 I(EXP-I) は 2 択の一般知識の 2 択問題を用いた情報カスケード実験である。実験 II(EXP-II) は赤 (a), 青 (b) の 2 色の玉を異なる割合で混ぜた 2 個の壺 A, B からランダムに選んだ壺 X が A, B どちらの壺だったかを推定する 2 択問題である。実験 I では被験者はノーヒントで回答した後、過去の被験者の回答 (2 択の各選択肢の選択者数) を参照して再度回答した。実験後、クリーニングしたデータに対してノーヒントでの正答率 $Z_0(T)$ が平均 50%, 60%, 70%, 80% で問題を分類してそれぞれのプールで解析を行う。実験 II では被験者は壺 X からランダムに選んだ玉の情報と過去の被験者の回答 (壺 A, B を選んだ被験者数) を参照して回答する。被験者の選んだ玉の色 a, b と壺 A, B が一致する事後確率 q_* は $q \in \{5/9, 6/9, 7/9\}$ である。サンプル数 I , 最初の被験者の選択が $X(1) = \{1, 0\}$ でのサンプル数 N_1, N_0 , 被験者数 (カスケード実験での被験者の列の長さ) T の情報は表 I にまとめてある。各実験およびデータ解析の詳細は論文または文献を参照のこと [Mori 12, Mori 13, Mori 14, Mori 15b, Mori 15c, Hino 15].

図 6 は各実験での $C(t)$ を t に対してプロットしたものである。実験 I ではノーヒントでの正答率 $Z_0(T)$ が 50%, 60% のサンプル、実験 II では自己情報 q が 5/9, 6/9 の場合、 $C(t)$ は有限の正の値のまわりでゆるぎようになる。これは $c > 0$ を示唆する。一方、実験 I で $Z_0(T)$ が 70%, 80% のサンプル、および実験 II の $q = 7/9$ の場

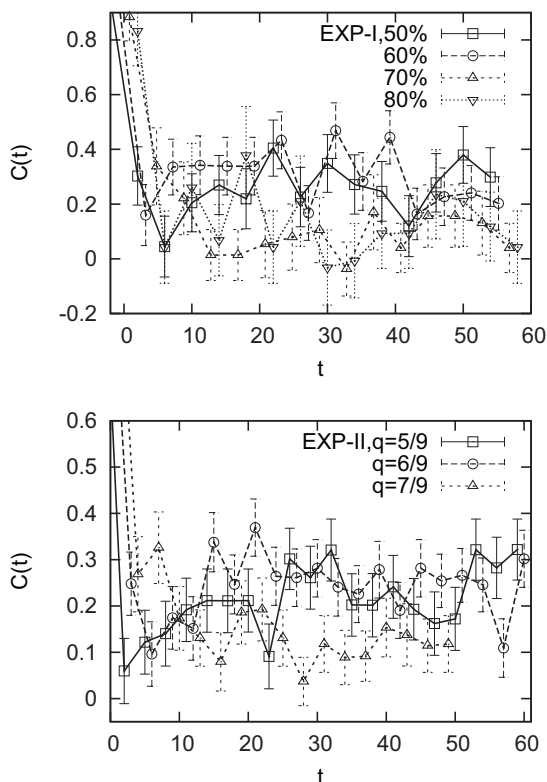


図 6 実験 I (上図), 実験 II (下図) の相関関数 $C(t)$ を t に対してプロット. データは $\Delta t = 2$ または 3 のみ, また, t 軸方向に ± 1 ないし ± 2 ずらしてプロットしている.

合, $C(t)$ は t とともに減衰する. ただし, 実験 I, II ともに $\Delta C(t)$ は大きく, 漸近形 (7) を仮定してフィットすることにより, c, l を評価することは難しいが, 実験 I の 50%, 60%, および実験 II の $q = 5/9, 6/9$ での c は $0.2 \sim 0.3$ 程度と考えられる.

図 7 は τ_t を t に対して両対数プロットしたものである. $c = 0$ のとき, $\tau_t \propto t^{l-1}$ と振る舞うため, プロットの傾きから l の値を推定できる. $C(t)$ のプロットから推測されたように, 実験 I の 50%, 60%, および実験 II の $q = 5/9, 6/9$ において τ_t は t とともに正の値に収束している. $\lim_{t \rightarrow \infty} \tau_t(t) = c$ より, その値が c の推定値である. 実験 I の上記の 2 ケースではそれぞれ $c = 0.31, 0.29$, 実験 II の $q = 5/9, 6/9$ でそれぞれ $c = 0.22, 0.26$ となった. 実験 I の 70%, 80% および実験 II の $q = 7/9$ では τ_t のプロットは右下がりの直線となっている. この直線の傾きから l を評価すると, 実験 I ではそれぞれ $l = 0.57, 0.68$, 実験 II では $l = 0.645$ となった.

図 8 は ξ_t を t に対してプロットしたものである. $c > 0$ の場合, ξ_t は $1/\sqrt{3}$ に, $c = 0$ の場合, $\sqrt{l/l+2}$ に収束する. l の値は τ_t の両対数プロットの傾きから評価したものをを用いて実線および破線, 点線でプロットしている. $C(t)$ および $\tau_t(t)$ のプロットで $c > 0$ と推測されたケースでは, 誤差の範囲内で $\xi_t(t)$ は $1/\sqrt{3}$ に収束している. また, $c = 0$ の場合, $\xi_t(t)$ は $\tau_t(t)$ のプロットで評価した l の値

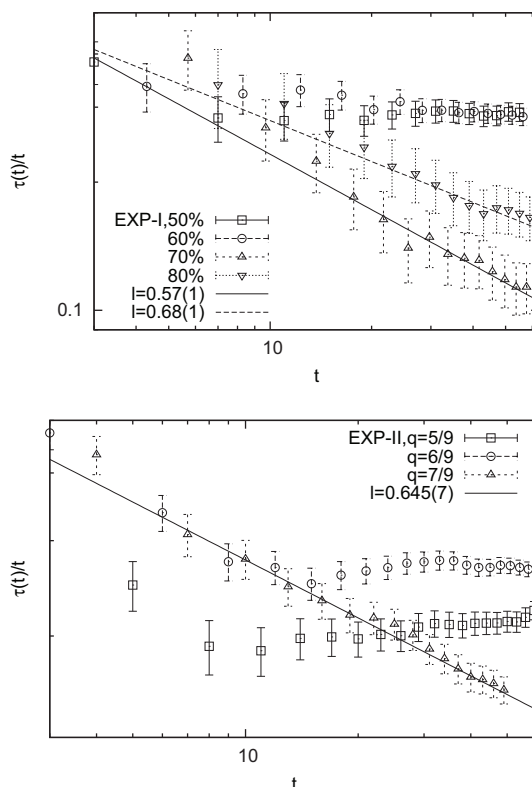


図 7 $\tau_t \equiv \tau(t)/t$ を t に対してプロット. 実線, および破線は a, l をパラメータとして $\tau_t = at^{l-1}$ でフィットした結果.

に対する $\sqrt{l/l+2}$ に誤差の範囲内で収束している.

以上の結果から, 実験 I のノーヒントでの正答率が 50%, 60%, および実験 II の自己情報の精度 $q = 5/9, 6/9$ の場合, $c > 0$ となり, 系は 2 個以上の安定状態を持つ相にある. 実験 I の 70%, 80%, および実験 II の $q = 7/9$ の場合, $c = 0$ であり, 系は 1 個の安定状態を持つ. クイズの難易度 $Z_0(T)$ および自己情報の精度 q をコントロールすることにより, 系は非線形ポリア壺の相転移を起こす.

4. ま と め

非線形ポリア壺の写像 $q(z)$ の変化による安定状態の個数の変化を相転移の観点から議論した. 非線形ポリア壺の相関関数 $C(t)$ が吸収状態相転移での $\rho(t)$ と類似の現象論的スケーリング則を満たすことから, 非線形ポリア壺の上記の変化は $C(t)$ の極限值 $c = \lim_{t \rightarrow \infty} C(t)$ を秩序変数とする非平衡相転移であることを述べた. また, $C(t)$ の漸近形 $C(t) \simeq c + c' \cdot t^{-1}$ に基づき, 実験データを用いて相転移を検証する方法について述べた. それを用いて 2 種類の情報カスケード実験のデータを解析し, 相転移を検証した. 被験者集団の持つ自己情報 (ノーヒントでの正答率やヒント) の精度が高い場合, 系は安定状態が 1 個の状態にあるが, 精度が低くなると安定状態が 2 個以上の状態に相転移した. 実験データのマイクロ解

◇ 参考文献 ◇

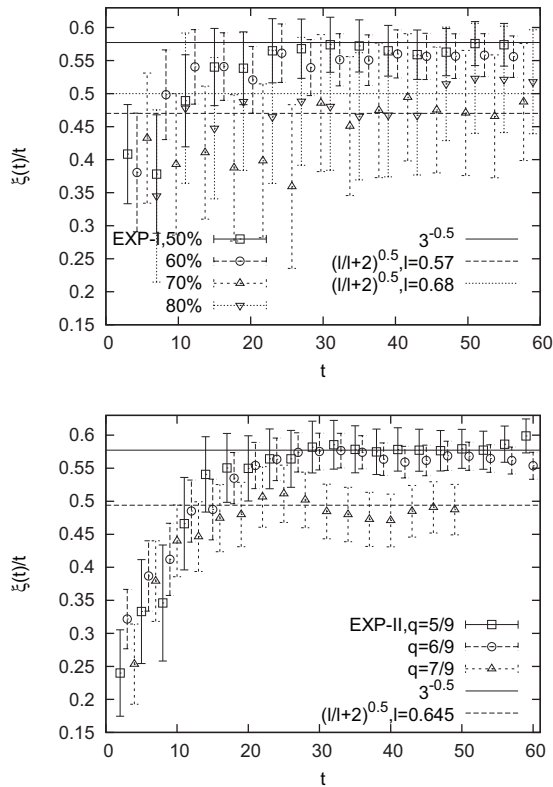


図 8 $\xi_t \equiv \xi(t)/t$ を t に対してプロット.

析からは、 $q(z)$ の安定固定点がそれぞれ 1 個と 2 個にあると分かっており、それらの結果と整合的である。

今後の問題は、非線形ポリア壺の写像 $q(z)$ のミクロな解釈・モデル化である。実験系経済学では、ベイズ統計やゲーム理論を用いてヒトの選択のモデル化が精密に行われ、実験データの説明にも用いられてきた。一方、我々はそうしたミクロなモデルからではなく、マクロに相転移が起きるのかどうか、という観点で実験データ駆動型のナイーブな $q(z)$ のモデルを採用し、非線形ポリア壺の理論で解析を行ってきた。こうした 2 つのアプローチを結ぶために実験データで推定した $q(z)$ を説明するモデルを構築する必要がある [Equiluz 15]。その際、過去の被験者の選択情報のうち、参照できる人数 r を固定した場合の $q(z)$ を測定し、 $q(z)$ の r 依存性を解析する必要があると考えられる。また、理論的には、デジタル型とイジング型の間で何が起きるのかも興味深い。 $q(z)$ の touch point での左微分が $1/2$ のとき、touch point が安定なのかどうか。また、参照人数 r をゆっくり増やしたとき平均場イジング模型の平衡相転移が確認できるのかどうか [Hisakado 15]。こうした問題の理解は非平衡相転移の理解を深めるであろう。

謝 辞

本研究は科研費 25610109 (挑戦的萌芽研究) の助成を受けた。

- [BHW 92] S.Bikhchandani and D.Hirshleifer and I.Welch: A Theory of Fads, Fashion, Custom, and Cultural Changes as Informational Cascades, *J. Polit. Econ.*, vol.100,992-1026(1992).
- [Anderson 97] L.R.Anderson and C.A.Holt: Information Cascades in the Laboratory, *Am. Econ. Rev.*, vol.87,847(1997).
- [Goeree 07] J.K.Goeree, T.R.Palfrey, B.W. Rogers and R.D.McKelvey: Self-Correcting Information Cascades, *Rev. Econ. Stud.*, vol.74,733-762(2007).
- [Arthur 90] W.B.Arthur: Positive feedbacks in the Economy, *Scientific American*, vol.262(1990)92-99.
- [Mori 12] S.Mori, M.Hisakado and T.Takahashi: Phase transition to two-peaks phase in an information cascade voting experiment, *Phys.Rev.E*, vol.86,026109-026118(2013).
- [Mori 13] S.Mori, M.Hisakado and T.Takahashi: Collective Adoption of MaxMin Strategy in an Information Cascade Voting Experiment, *J.Phys.Soc.Jpn.*, vol.82,0840004-0840013(2013).
- [Hisakado 11] M.Hisakado and S.Mori: Digital Herders and Phase transition in a Voting model, *J.Phys.A.*, vol.44,275204-275220(2011).
- [Hisakado 12] M.Hisakado and S.Mori: Two kinds of Phase transitions in a Voting model, *J.Phys.A*, vol.45(2012)345002-345016.
- [Hill 80] B. Hill and D. Lane and W. Sudderth: A strong law for some generalized urn processes, *Ann. Prob.*, vol.8(1980)214-226.
- [Mori 15] S.Mori and M.Hisakado: Finite-size scaling analysis of binary stochastic processes and universality classes of information cascade phase transition, *J. Phys. Soc. Jpn.*, vol.84,054001-054013(2015).
- [Mori 15a] S.Mori and M. Hisakado: Correlation function for generalized Polya urns: Finite-size scaling analysis, arXiv:1501.00764.
- [Mori 15b] S.Mori, M. Hino, M. Hisakado and T.Takahashi: Detection of non-self-correcting nature in information cascade, accepted in ECCS'14 proceedings, arXiv:1507.07265.
- [Hino 15] M. Hino, M. Hisakado, T. Takahashi and S.Mori: Phase transition of generalized Pólya urn in Information cascade experiment, arXiv:1507.07269.
- [Pemantle 91] R.Pemantle: When are touchpoints limits for generalized Pólya urns?, *Proc. Amer. Math. Soc.*, vol.113(1991)235-243.
- [Hinrichsen 00] Hinrichsen: Nonequilibrium Critical Phenomena and Phase Transitions into Absorbing States, *Adv. Phys.*, vol.49(2000)815.
- [Mori 14] 守, 入江, 久門, 高橋: 2 択のクイズを用いた情報カスケード実験: 実験の手順とデータアーカイブについて, *統計数理研究所共同研究レポート 311 経済物理とその周辺 (10)*, 1-12(2014).
- [Mori 15c] 守, 入江, 久門, 高橋: 2 択のクイズを用いた情報カスケード実験: データ解析 (統計数理研究所共同研究レポート 332 経済物理とその周辺 (11)), 54-65(2015).
- [Equiluz 15] V. M. Equiluz and N. Masuda and J. Fernández-Gracia: Bayesian decision making un human collectives with binary choices, *PLoS One* 10(2015)e0121332-e0121345.
- [Hisakado 15] M. Hisakado and S.Mori: Information cascade, Kirman's ant colony model, and kinetic Ising model, *Physica A*. vol.417(2015)63-75.