

確率的順位付けの数理モデル

服部 哲 弥

1 強度が位置依存性を持つ確率的順位付けモデル

本稿で扱うのは、確率的順位付けモデルと呼ぶ、粒子たちの順位（位置）の時間発展についての確率微分方程式系（粒子系）と、この粒子系の流体力学極限と呼ぶ、粒子数の大きい極限である。

予告を加えると、下記の方程式 (1) で決まる粒子系のランダムで従属性のある時間発展は、(24) のように、粒子数 N が大きいときランダムだが独立な確率過程に近い（収束する）。ここで、(1) にない量 μ_t は、ランダムな分布（経験分布）(4) の極限だが非ランダムな分布であり、流体力学極限と呼んだとおり、流体成分の空間分布の時間発展と解釈できる。以上の結果を本節にまとめ、第 2 節で簡単な場合の結果と実社会での具体例、第 3 節で主定理の証明の構造、を文献を読む際の手がかりになるように紹介する。

1.1 確率的順位付けモデル

$T > 0$ を以後固定して有界な $1 + 1$ 次元時空を考え、位置と時刻を $(y, t) \in [0, 1] \times [0, T]$ で表す。自然数 N を固定することに本稿で扱う方程式に現れる $i = 1, \dots, N$ と $t \in [0, T]$ に対する（標準の確率空間 (Ω, \mathcal{F}, P) 上の）確率変数 $Y_i^{(N)}(t) : \Omega \rightarrow [0, 1]$ を時刻 t での粒子 i の位置または相対順位と見る。いつものように以後サンプル $\omega \in \Omega$ への言及を概ね略す。

第 1 象限 $\mathbb{R}_+ \times \mathbb{R}_+$ 上の単位強度のポワソン乱測度 $\nu_i^{(N)}$ の独立な列 $i = 1, \dots, N$ によって確率が方程式に入る。ここで、第 1 象限 $\mathbb{R}_+ \times \mathbb{R}_+$ 上の単位強度のポワソン乱測度 ν とは、(第 1 象限の 2 次元ルベグ可測部分) 集合 $A \subset \mathbb{R}_+^2$ に対して $U = \nu_i^{(N)}(A)$ は平均が面積 $(E[U([a, b] \times [c, d])]) = (b - a)(d - c)$ のポワソン分布 $(P[U = k] = e^{-E[U]} \frac{E[U]^k}{k!})$ に従う確率変数で、第 1 象限内の排反 $(A \cap B = \emptyset)$ な集合に対して $U(A)$ と $U(B)$ が独立なものを言う。

各 N ごとに、各粒子 $i = 1, \dots, N$ について、 $\nu_i^{(N)}$ の $Y_i^{(N)}(t)$ への影響の強さを与える非負実数値連続関数 $w_i(y, t)$ （強度）を選んで固定し、初期順位 $Y_i^{(N)}(0) = y_i^{(N)}$ も固定する ($y_i^{(N)} = \frac{i-1}{N}$) と、下記の方程式 (1) によって粒子系の確率的な時間発展が一意に決まる。これが確率的順位付けモデルである。なお、以下、集合 A 上で 1 で補集合 A^c 上で 0 なる関数を 1_A と書き、また、時刻 s の関数 $Y_i^{(N)}(s)$ において s についての左極限の値を $Y_i^{(N)}(s-)$ と書く。

$$\begin{aligned}
 Y_i^{(N)}(t) &= y_i^{(N)} \\
 &+ \frac{1}{N} \sum_{j=1}^N \int_{s \in (0, t]} \int_{\xi \in \mathbb{R}_+} \mathbf{1}_{Y_j^{(N)}(s-) > Y_i^{(N)}(s-)} \mathbf{1}_{\xi \in [0, w_j(Y_j^{(N)}(s-), s))} \nu_j^{(N)}(d\xi \times ds) \\
 &- \int_{s \in (0, t]} \int_{\xi \in \mathbb{R}_+} Y_i^{(N)}(s-) \mathbf{1}_{\xi \in [0, w_i(Y_i^{(N)}(s-), s))} \nu_i^{(N)}(d\xi \times ds), \\
 &i = 1, 2, \dots, N, \quad t \geq 0.
 \end{aligned} \tag{1}$$

見た目を少し短くするために、

$$\tilde{\nu}_i^{(N)}(t) = \int_{s \in (0, t]} \int_{\xi \in \mathbb{R}_+} \mathbf{1}_{\xi \in [0, w_i(Y_i^{(N)}(s-), s))} \nu_i^{(N)}(d\xi \times ds) \quad (2)$$

とにおいて (1) に代入すると,

$$Y_i^{(N)}(t) = y_i^{(N)} + \frac{1}{N} \sum_{j=1}^N \int_0^t \mathbf{1}_{Y_j^{(N)}(s-) > Y_i^{(N)}(s-)} \tilde{\nu}_j^{(N)}(ds) - \int_0^t Y_i^{(N)}(s-) \tilde{\nu}_i^{(N)}(ds) \quad (3)$$

と書ける.

$w_i(y, s)$ が定数ならば (2) の $\tilde{\nu}_i^{(N)}$ は, $\tilde{\nu}_i^{(N)}(t) = \nu_i^{(N)}([0, w_i] \times (0, t])$ なので, $\tilde{\nu}_i^{(N)}$ は強度 w_i のポワソン過程である. $w_i(y, s)$ が y について定数ならば時刻 s に依存しても $\tilde{\nu}_i^{(N)}$ はやはり独立増分な確率過程 (非一様ポワソン過程) である. このとき粒子 i の位置 $Y_i^{(N)}$ は, 他の j と独立にポワソン到着時刻ごとに 0 になる (3) の右辺第 3 項の効果と, 第 2 項による自分より下位の j が先頭に跳んだ結果追い越されて $\frac{1}{N}$ 単位で順位を下げる効果だけで時間変化する. ここで $Y_j^{(N)}(t) > Y_i^{(N)}(t)$ であることを (時刻 t で) j が i の下位にあると表現するが, 模型の名前どおり, 位置 $Y_i^{(N)}$ を相対順位と見ている. たとえば位置 0 は 1 位を意味する. このアルゴリズムは先頭に跳ぶ (move-to-front) 規則として古くから知られている [46] が, w_i たちが位置に依存しない場合に (3) をあらわに書いたのは永幡幸生さん [40], 位置依存性がある場合に $\nu_i^{(N)}$ を導入して, 追加した仮想的な次元方向に $w_i^{(N)}$ を取って (1) を書いたのは楠岡誠一郎さん [30] である.

先頭に跳ぶ規則 (3) において, 注目する粒子への他粒子からの直接の影響 (確率変数の従属性) は下位の粒子たちの総和だけであることから, 時間発展と極限を考える上で自然な量は, δ_c を 1 点 c に集中した単位分布として, 位置強度結合経験分布

$$\mu_i^{(N)} = \frac{1}{N} \sum_{i=1}^N \delta_{(w_i, Y_i^{(N)}(t))} \quad (4)$$

と, (3) で先頭に跳ぶ第 3 項を落として $\gamma = (y_0, t_0) \in [0, 1] \times [0, T]$ と t の関数とした

$$Y_C^{(N)}(\gamma, t) = y_0 + \frac{1}{N} \sum_{j=1}^N \int_{t_0}^t \mathbf{1}_{Y_j^{(N)}(s-) > Y_C^{(N)}(\gamma, s-)} \tilde{\nu}_j^{(N)}(ds), \quad t \geq t_0, \quad (5)$$

である. (5) は粒子たちに挟まれて流れに従って下流に移動するが先頭には跳ばない時間発展なので, (極限流体の時間発展を定める方程式の) 特性曲線の離散近似である. $\mu^{(N)}$ は強度関数 w_i たちの集合 W と位置 $[0, 1]$ の直積空間上の確率測度 (分布) に値を取る確率過程である, $\mu^{(N)}$ はばらつく (サンプル $\omega \in \Omega$ の関数である) が, 初期値 $\mu_0^{(N)}$ が $N \rightarrow \infty$ でサンプルによらない分布 μ_0 に収束すれば, サンプルによらない分布 μ_t の時間発展に概収束する, ということが主定理である.

本稿では, 粒子系の具体的なイメージは (3) と (4) の間でも言及したとおりランキング (順位付け), 流体力学極限を取った系のイメージは蒸発による (湖で行き止まりの淀んだ川や樋に海水を流して蒸発で塩を作る古典的な仕組みのような) 1 方向への流れを考える. (追い越しのない 1 次元の流れという点だけを考えると, 毛細管の中の血球のほうは誤解が少ないかもしれないが.) 流体では先頭への跳びは蒸発した流体の各成分量を保存するように上流から供給される大気循環の模式を意味する.

1.2 流体力学極限

現実との関係について続きを説明する前に, 数学的な話が紛れないように, 主定理を書き切ってお

く．極限について発見した(それなりに)あらわな形を記述するために, 必要な準備が少し長くなる．背景や簡単な場合の表式の例に興味があれば, 本小節を後に回して第 2 節を先に読んでいただいて差し支えない．(逆に背景等よりも主定理の証明の組み立てに興味がある場合は, この小節を終えた後, 第 2 節を飛ばして第 3 節に進んで差し支えない.)

$\gamma = (y_0, t_0) \in [0, 1] \times [0, T], t \geq t_0, w \in W$ に対して

$$\varphi^{(N)}(dw, \gamma, t) = \mu_t^{(N)}(dw \times [Y_C^{(N)}(\gamma, t), 1]) \quad (6)$$

とにおいて, $\mu_t^{(N)}$ の位置についての分布関数である測度値確率過程 $\varphi^{(N)}$ を定義する．通常の分布関数と違って 0 からの区間ではなく下流側を取っていることと端点が特性曲線の位置であることは先頭に跳ぶ規則用の調整である．粒子が間隔 N^{-1} で並んでいて, 粒子の種類(強度)を見なければ区間 $(a, 1)$ に $[N(1-a)]$ (大括弧は整数部分) 個の粒子があることに注意すると (6) から

$$Y_C^{(N)}(\gamma, t) = y_0 + \frac{[N(1-y_0)]}{N} - \varphi^{(N)}(W, \gamma, t), \quad (7)$$

を得るので, 特性曲線の精密化と見ることもできる自然な対象である．

強度 w_i たちが位置依存性を持たなければ $\varphi^{(N)}$ に注目するだけで本質的に独立確率変数列の大数の法則に帰着し, 極限はポワソン過程の言葉で書けるが, 位置依存性があれば従属性の影響で極限を記述する確率過程は独立増分性を失い, 極限の流れも複雑になって流れの集合上の写像の固定点として得られる．以上を記述する記号を用意する．時空 $[0, 1] \times [0, T]$ の一方向への流れの発展方程式の初期 ($t = 0$)・境界点 ($y = 0$) の集合を

$$\Gamma = \Gamma_b \cup \Gamma_i, \quad \Gamma_b = \{(0, s) \mid 0 \leq s \leq T\}, \quad \Gamma_i = \{(z, 0) \mid 0 \leq z \leq 1\}, \quad (8)$$

と置き, そのうち時刻 t 時点での流れに影響する部分を

$$\Gamma_t = \{(z, t_0) \in \Gamma \mid t_0 \leq t\} = \Gamma_i \cup \{(0, t_0) \in \Gamma_b \mid 0 \leq t_0 \leq t\}, \quad (9)$$

と置いて, 時刻 t との整合的な組を

$$\Delta_T = \{(\gamma, t) \in \Gamma_T \times [0, T] \mid \gamma \in \Gamma_t\}. \quad (10)$$

と置く．この記号を用いて流れの集合を

$$\Theta_T = \{\theta: \Delta_T \rightarrow [0, 1] \mid \text{連続}, \theta((y_0, t_0), t_0) = y_0, (y_0, t_0) \in \Gamma_T, \\ \text{全射}, \text{各 } t \text{ で } \gamma \text{ について非増加}, \text{各 } \gamma \text{ で } t \text{ について非減少}\}, \quad (11)$$

と置く． γ について増加減少と書いたが, Γ 上の大小関係は, (長方形の 2 辺を素直に延ばして,)

$$s \leq t, z \leq y \Leftrightarrow (0, T) \succeq (0, t) \succeq (0, s) \succeq (0, 0) \succeq (z, 0) \succeq (y, 0) \succeq (1, 0), \quad (12)$$

で定義する．

時空の関数である強度 $w \in W$ の位置変数に流れ $\theta \in \Theta_T$ を代入し ($s = 0$ の場合はさらに初期位置 $z \in [0, 1]$ を選んで),

$$\tilde{w}_{\theta, w, z}(s, t) = \begin{cases} w(\theta((z, 0), t), t), & s = 0, \\ w(\theta((0, s), t), t), & s > 0, \end{cases} \quad (13)$$



と置いて、流れの連結成分（流線）の起点時刻と現在時刻の関数と見た $\tilde{w}_{\theta,w,z}(s,t)$ を強度とする、直前の到着時刻に強度が依存する点過程を $\tilde{\nu}_{\theta,w,z}$ とおく。ここで、 $w(s,t)$ を強度とする、直前の到着時刻に強度が依存する点過程（非減少右連続非負整数値初期値 0 の確率過程） $\tilde{\nu}$ とは、 $\tau_0 = 0$ として、各正整数点 k への到着時刻 $\tau_k = \inf\{t \geq 0 \mid \tilde{\nu}(t) \geq k\}$ の、直前の到着時刻 τ_{k-1} についての条件付き確率が $t \geq \tau_{k-1}$ に対して

$$P[t < \tau_k \mid \mathcal{F}_{\tau_{k-1}}] = \exp\left(-\int_{\tau_{k-1}}^t w(\tau_{k-1}, u) du\right) \quad (14)$$

で与えられるものをそう名付けた [25, 26]。 $\tilde{\nu}_{\theta,w,z}$ を用いて、流れ $\theta \in \Theta_T$ と（第 1.1 節の終わり近くで主定理の仮定の 1 つとして予告した）位置強度結合経験分布の初期値の極限 μ_0 に対して、

$$\begin{aligned} \varphi_\theta(dw, \gamma, t) &= \int_{z \in [y_0, 1]} P[\tilde{\nu}_{\theta,w,z}(t) = \tilde{\nu}_{\theta,w,z}(t_0)] \mu_0(dw \times dz), \\ \gamma &= (y_0, t_0) \in \Gamma_t, (\gamma, t) \in \Delta_T, \end{aligned} \quad (15)$$

で φ_θ を定義する。この φ_θ が分布関数 (6) の極限になるように θ を適切に選ぶ。(7) の連続極限版によって φ_θ は特性曲線 $y_C = y_C(\gamma, t)$ を導くことを見越すと、その起点 γ について曲線たちを集めると流れが定義できるので、 $\mathcal{G}(\theta)(\gamma, t) = 1 - \varphi_\theta(W, \gamma, t)$ で流れの集合上の写像 $\mathcal{G} : \Theta_T \rightarrow \Theta_T$ を定義できる。連続極限を与える θ は \mathcal{G} の固定点 $y_C = \mathcal{G}(y_C)$ である。ここまで強度 $w \in W$ のなめらかさについて仮定しなかったが、次の固定点 y_C の唯一存在定理と主定理（収束定理）の両方に十分な条件を置く。以下 $W = C^{1,0}([0, 1] \times [0, T])$ 、すなわち、連続関数 w_i たちは位置変数に関する導関数が連続であることも仮定する。位置依存性については、おとなしくさらに、導関数がある界であること

$$C_W := \sup_{w \in W} \sup_{(y,t) \in [0,1] \times [0,T]} \left| \frac{\partial w}{\partial y}(y, t) \right| < \infty \quad (16)$$

を仮定する。

定理 1 ([25, Theorem 9, (89)]) $\theta \in \Theta_T$ についての関数方程式

$$\theta(\gamma, t) = \mathcal{G}(\theta)(\gamma, t) := 1 - \varphi_\theta(W, \gamma, t), \quad \gamma = (y_0, t_0) \in \Gamma, (\gamma, t) \in \Delta_T, \quad (17)$$

を満たす $\theta = y_C \in \Theta_T$ がただ一つ存在する。 ◇

分布関数と分布が 1 : 1 であるのと同様に、

$$\varphi_{y_C}(dw, (y_0, t_0), t) = \mu_t(dw \times [y_C((y_0, t_0), t), 1]), \quad ((y_0, t_0), t) \in \Delta_T, \quad (18)$$

を満たす分布 μ_t がただ 1 つ存在し、定理 1 の y_C は

$$y_C(\gamma, t) = y_0 + \int_{t_0}^t \int_{W \times [y_C(\gamma, s), 1]} w(z, s) \mu_s(dw \times dz) ds \quad (19)$$

を満たすこともわかる [25, Theorem 1 の下]。定理 1 では固定点と書いたが、(16) の下で \mathcal{G} は縮小写像 [25] なので、解 y_C は（任意の、すなわち、単純な）流れ $\theta_0 \in \Theta_T$ から $y_C = \lim_{n \rightarrow \infty} \mathcal{G}^n(\theta_0)$ で得られる。

最大値ノルムに基づく標準の距離によって、 $C^{1,0}([0, 1] \times [0, T])$ は完備可分距離空間になり [3, Example 26.2]、本稿の測度たちはボレル測度として分布の列の弱収束が定義できる [5]。元の確率的順



位付けモデルの初期分布について $N \rightarrow \infty$ で $\mu_0^{(N)} \rightarrow \mu_0$ のときに $\mu_t^{(N)} \rightarrow \mu_t$ となることが主定理である。細かく書くと、 $t = 0$ では

$$\begin{aligned} & \text{任意の一樣有界同程度連続な集合 } H \subset C^0(W \times [0, 1]; \mathbb{R}) \text{ に対して} \\ & \exists \delta \in (0, \frac{1}{2}), \exists C > 0; (\forall N \in \mathbb{N}) (\forall \tilde{h} \in H) (\forall y \in [0, 1]) \\ & \left| \int_{W \times [y, 1]} \tilde{h}(w, z) \mu_0^{(N)}(dw \times dz) - \int_{W \times [y, 1]} \tilde{h}(w, z) \mu_0(dw \times dz) \right| \leq \frac{C}{N\delta}, \end{aligned} \quad (20)$$

を以下仮定する。 λ を μ_0 の W への周辺分布 $\lambda(dw) = \mu_0(dw \times [0, 1])$ とする。(ちなみに、強度は粒子ごとに固定するので N, t に関係なく $\lambda^{(N)} = \frac{1}{N} \sum_{i=1}^N \delta_{w_i}$ および $\lambda(dw) = \mu_t(dw \times [0, 1])$ を得る。また、位置は相対順位で等間隔で密度変化がないので $\mu_t(W \times dy) = dy$ 等となる。) $\lambda^{(N)}$ と λ については、 $\|w\|_T$ を最大値ノルムとして

$$\lim_{N \rightarrow \infty} \int_W \|w\|_T \lambda^{(N)}(dw) = \int_W \|w\|_T \lambda(dw) < \infty \quad (21)$$

を仮定する。 $\lambda^{(N)}$ と λ の定義から仮定 (20) は $\lambda^{(N)} \rightarrow \lambda, N \rightarrow \infty$ を意味する。

定理 2 (主定理, [28, Theorem 2]) 仮定 (16), (20), (21) の下で、確率 1 で $N \rightarrow \infty$ のとき (4) の $\mu_t^{(N)}$ は t について一様に (18) の μ_t に弱収束する。すなわち、任意の $y \in [0, 1]$ と有界連続関数 $h: W \rightarrow \mathbb{R}$ に対して

$$\lim_{N \rightarrow \infty} \sup_{t \in [0, T]} \left| \int_W h(w) \mu_t^{(N)}(dw \times [y, 1]) - \int_W h(w) \mu_t(dw \times [y, 1]) \right| = 0, \text{ a.e.}, \quad (22)$$

が成り立つ。

さらに、カオスの伝搬 (名札付き部分有限粒子系の独立確率過程への収束, propagation of chaos) が成り立つ。すなわち、正整数 L と $y_i \in [0, 1], i = 1, 2, \dots, L$, に対して

$$i = 1, 2, \dots, L, \text{ について } \nu_i^{(N)} = \nu_i, N \in \mathbb{N}, \text{ および } \lim_{N \rightarrow \infty} y_i^{(N)} = y_i, \quad (23)$$

が成り立つとき、名札付き粒子系 $(Y_1^{(N)}(t), Y_2^{(N)}(t), \dots, Y_L^{(N)}(t))$ は確率 1 で $N \rightarrow \infty$ のとき $t \in [0, T]$ について一様に $(Y_1(t), Y_2(t), \dots, Y_L(t))$ に収束する。ここで、各 $i = 1, 2, \dots, L$ に対して Y_i は

$$\begin{aligned} Y_i(t) &= y_i + \int_{s \in (0, t]} \int_{(w, z) \in W \times [Y_i(s-), 1]} w(z, s) \mu_s(dw \times dz) ds \\ &\quad - \int_{s \in (0, t]} \int_{\xi \in \mathbb{R}_+} Y_i(s-) \mathbf{1}_{\xi \in [0, w_i(Y_i(s-), s))} \nu_i(d\xi \times ds), \\ & i = 1, 2, \dots, L, t \in [0, T], \end{aligned} \quad (24)$$

の唯一解である。 \diamond

主定理は確率 1 のサンプル $\omega \in \Omega$ ごとの収束という意味で概収束であり、サンプル間で有限粒子数の位置強度結合経験分布 $\mu_t^{(N)}$ がばらつく一方、極限 μ_t はサンプルによらないという意味で、広くくりでは (従属測度値確率変数列の) 大数の強法則である。ただし、独立性は N を固定するごとに粒子間で ($\nu_i^{(N)}$ が異なる i の間で独立という形で) のみ仮定している点で、通常のランダムウォークの

大数の強法則よりも強い大数の完全法則 [33, 13, 14] に対応する [29]. 第 3 節で紹介するように, その強い評価が証明の途中段階で中間模型の収束において用いられる [25].

第 2 節で順次紹介するように, w が位置変数について定数ならば証明は相対的に簡単であり, 簡単であるにも関わらず動機となった応用に関していくつかのことを説明できる.

2 強度が位置依存性を持たない確率的順位付け模型

2.1 Amazon ランキング

小売業の構造を大きく変えた巨大なネット小売業の印象がある Amazon だが, 元々はネット書店であった. 和数学書を心に決めて, 日本法人 Amazon.co.jp の検索欄で「本」のカテゴリーを選び, 検索して当該書のウェブページを開くと, 他の本の宣伝のやや下「登録情報」のところにランキングという数字がある.(和書としたのは Amazon ランキングが和洋でわかれているため, 数学書としたのは読者の関心を引こうとしただけで実際には数学書に限らずどの本を選んでも同じ現象が見られることが本記事の内容に含まれている.) 大流行の著名本は別格として, 大半の本の数十万位という順位は自著でのみ気になる数字だが, 1 時間に 1 度しか更新しないこともあって, 思い出したように見に行くと, たまに桁を間違えたかと思う順位がついて謎めく. 「ヘルプ」の説明の文言はときおり更新されるが, たとえば「最新販売数と累計販売数を反映」して決まるといった, 誰もが当然視する内容の漠然とした追認に終わる. 最新販売数, 言い換えると, 流行度を順位づけているというが, まれにしか売れない大多数の本の流行度とはそもそも何か, どこから「最新」なのか, 遠過去と近過去の重みはどうするか, など実用的にアプローチしようとする謎が解けそうにない. 発想を転換して最近の人気や流行を反映するもっとも単純な順位付けを目指す, 最後に売れたものを 1 位とする「先頭に跳ぶ規則」に行き着く [23]. 先頭に跳ぶ規則に加えて, 本のネット注文(書店側から見れば, 販売)は不特定

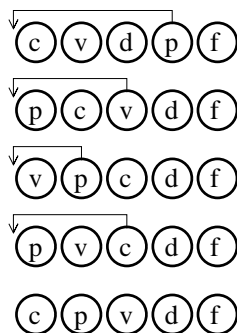


図 1: 「先頭に跳ぶ規則」は流行度の順位付けのもっとも簡単な定義(本文では書名 c, d, f, p, v を味気なく $i = 1, \dots, N$ と書く.)

多数の読者の独立ランダムな行動として, 点過程として広く用いられるポワソン過程で単純化すれば, 確率的順位付け模型を得る [20, 21, 22, 19, 24, 23, 40, 41, 30, 25, 26, 27, 29, 28]. w_i が位置依存性を持つと第 1.2 節のように結果が複雑なので, ここではまず研究の出発点となったもっとも単純な場合, すなわち, w_i が定数の場合を扱う. このとき $W = C^{1,0}([0, 1] \times [0, T])$ は $W = \mathbb{R}_+$ と同一視できて, たとえば, 強度の分布 $\lambda(dw) = \mu_0(dw \times [0, 1])$, すなわち, 流体成分の分布, 等は \mathbb{R}_+ 上

の, いわゆる 1 次元分布と同一視できる. しかも, (2) の $\tilde{\nu}_i^{(N)}$ が強度 w_i のポワソン過程で i について独立だから, 極限も含めて確率は通常のポワソン過程で書けるので, 下流側で指数分布に従って蒸発した量の分だけ流れが進むという意味の, 以下のようなわかりやすい結果が成り立つ [20, 21].

まず, 最上流を起点 ($\gamma = (y_0, t_0) = (0, t_0)$) とする (5) の特性曲線の極限 $y_C(\gamma, t)$ は

$$y_C(\gamma, t) = 1 - \int_0^\infty e^{-w(t-t_0)} \lambda(dw), \quad t \geq t_0, \quad \gamma = (0, t_0), \quad (25)$$

と, 強度の分布 λ を用いて具体的に書ける. また, 増加関数 $y = y_C((0, 0), t)$ の逆関数を $t_C : [0, 1) \rightarrow \mathbb{R}_+$ と置くと, (25) から $t_C(y_C(\gamma, t)) = t - t_0$ および $y = 1 - \int_0^\infty e^{-wt_C(y)} \lambda(dw)$ であって,

$$\mu_t(dw \times [y, 1)) = \mu(dw \times [y, 1)) = e^{-wt_C(y)} \lambda(dw), \quad (26)$$

あるいは, 微分 $t'_C(y_C) y'_C = 1$ に (25) を使った後 $y_C = y$ と $t - t_0 = t_C(y)$ で書き直して

$$\mu(dw \times dy) = \frac{we^{-wt_C(y)}}{\int_0^\infty ve^{-vt_C(y)} \lambda(dv)} \lambda(dw) dy$$

が定常分布であることもわかる. (第 1.2 節も既読の場合は, 直前の到着時刻に強度が依存する点過程 $\tilde{\nu}_{\theta, w, z}$ が, 強度が一定ならポワソン過程になることから, (15) はポワソン過程で時間 $[t_0, t)$ の到着が 0 回の確率 (指数分布) で書いて, 定理 1 を用いることなく独立確率変数列の大数の法則の直接適用で極限が取れて特性曲線 $y_C = \theta$ について (25) を得る. (26) も (15) と (18) から得る.) 粒子 i の軌道の $N \rightarrow \infty$ 極限 (定理 2 の (24)) は, 強度 w_i のポワソン過程の k 回目の到着時刻を τ_k とおくと

$$Y_i(t) = y_C((0, \tau_k), t), \quad \tau_k \leq t < \tau_{k+1}, \quad k = 1, 2, \dots, \quad (27)$$

となる.

以上から, 1 部売れた本は 1 位という先頭に跳ぶ規則を, そのまま当てはめる暴挙が予言する Amazon ランキングの時間変化 $X_i^{(N)}(t)$ は, 定義 (1) では規格化して本 i が時刻 t で 1 位であることを $Y_i^{(N)}(t) = 0$, 最下位を $1 - \frac{1}{N}$ としたので, $X_i^{(N)}(t) = NY_i^{(N)}(t) + 1$ で与えられる. Amazon.co.jp の本は研究開始当時に N が百万 (10^6) の程度, 現時点でその約 2 倍と推測されるので, $N \rightarrow \infty$ の極限の結果が良い近似であろう. 本 i の累積 k 冊目の注文時刻を τ_k として, Amazon.co.jp における書籍の注文頻度の分布を λ とすると, (27) と (25) から, $\tau_k \leq t < \tau_{k+1}$ において

$$X_i^{(N)}(t) = NY_i(t) + o(N) = N \int_0^\infty (1 - e^{-w(t-\tau_k)}) \lambda(dw) + o(N) \quad (28)$$

を得る. 最後に本 i が売れた時刻 τ_k を時刻の起点に取り直せば (τ_k で条件付ければ) 右辺がバラツキのない量なのは, 多数の刊行書の売り上げのバラツキの相殺という, 大数の法則であり, 多数の本の売上の総計で相対的に順位を落とすので, どの本に注目しても (i によらず) 同形の曲線に従う.

曲線の形状の前に, 注文が入った本が 1 位になるという先頭に跳ぶ規則の顕著な特徴が実際に見えることを確認する必要がある. Amazon はランキングのアルゴリズムを自由に決められるので, たとえば流行はほとんど考慮せず累積販売数で順位を付けても, 「最新販売数と累計販売数を反映」という

漠然たる説明に矛盾しない．バグやシステム更新が特徴を見えなくする可能性もある．

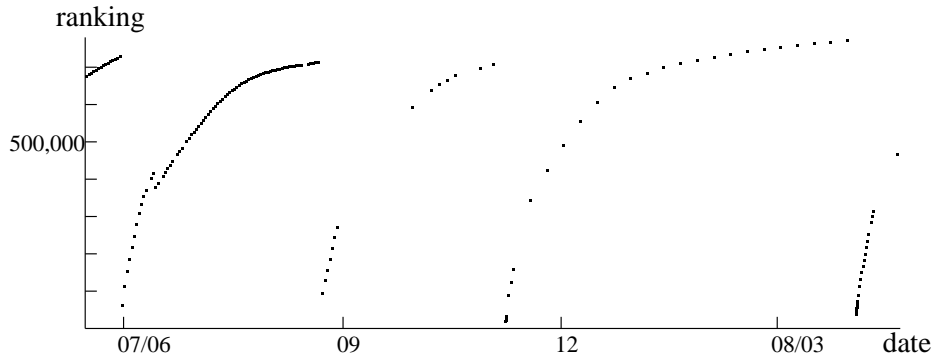


図 2: ランキングの時間変化の観測値．横軸は時刻 t ，全体で約 1 年，縦軸は X_i ，下が小さな値すなわち上位．

図 2 は筆者が初めて観測し記録した Amazon ランキングの時間発展である [22]．帆船の帆のような，ほぼなめらかな上に凸な増加曲線に並ぶ各点が観測値，最初の帆の密な点列は毎日定時の観測，最後の帆は週 1 度の観測である．順位が改善するときはそれまで落としていた順位が一気に横軸（1 位）付近まで跳ぶ．順位の一気の改善がその本の注文に対応することは，めったに売れない専門書を注文して 2 時間ほど順位の変化に注目すればわかる．桁の見間違いと見えていた良い順位は，本が売れた直後に見た場合の結果で，先頭に跳ぶ規則が Amazon ランキングの悪くない理想化であることの証であった．

先頭に跳ぶ規則は単純なだけに種々の別名で呼ばれる．文献上 [46] に遡ることは既に引用したが，同書には本の山から取り出して読んだら一番上に載せる「積ん読」を正当化する段落がある．同じアルゴリズムを「超整理法」と名付けた本もあった．窓口で処理されず諦めて離脱することだけで進む待ち行列と思えるし，時間反転すれば top-to-random shuffling でもある．Move-to-front という命名は 1970 年代 [39, 31, 9, 38, 37] のようである．その後しばらくの膨大な文献 [42, 15, 6, 11, 7, 43, 16, 18, 17, 35, 44, 36, 2, 32] は，テープのような順番に記録する古典的な媒体だけでなく，値段と速さが相反する記憶媒体系の計算実行時の様子に応じた使い分け（データの動的配置）のアルゴリズムの 1 つ (least-recently-used (LRU) caching) の理論的研究という位置づけのようである．（もっとも，現代では，OS カーネルの仮想記憶管理実装では前処理時間が大き過ぎて LRU はそのままでは使えないという話を，偶然拙著 [23] を見つけて連絡を下さった，筑波大の情報工学研究科の加藤和彦先生から最近伺った．高速化のぎりぎりのところを競うので，データ（ページ）が使われる度にそれより最近使われたページの順位を付け直すアルゴリズムは，さすがに単純化のしすぎのようである．）

2.2 Amazon はロングテールに非ず

確率的順位付け模型という数学的な単純さを追求したモデルが現実のランキングデータの特徴と整合していることが実測によってわかったので，もう一步踏み込んで図 2 の実際の Amazon ランキングと理論式 (28) を比べる．比べるには各書籍の平均売上の分布 λ が必要である．企業の種々の詳しいデータは秘匿されているので，分布の概形（分布族）を指定して，パラメータ（母数）をデータから統計的に当てはめる必要がある．（年齢などの属性を適切に揃えたときの）身長分布のように平均

とそのまわりの揺らぎの描像で語られる場合は正規分布族，所得などのように格差が議論となるときはべき分布（一般化パレート分布，離散版 λ_N では一般化ジップの法則）族が広く使われる．専門書と流行書を思い浮かべるまでもなくここでは後者を選んで，

$$\lambda([w, \infty)) = \left(\frac{a}{\max\{w, a\}} \right)^b \quad (29)$$

と置く． a と b は正定数である．(29) を (28) に代入すると，時刻 t 時点で最後に売れた時刻を τ_k として，

$$X_i^{(N)}(t) \sim N - Ne^{-a(t-\tau_k)} + N(a(t-\tau_k))^b \Gamma(1-b, a(t-\tau_k)) \quad (30)$$

と，不完全ガンマ関数 $\Gamma(z, p) = \int_p^\infty e^{-x} x^{z-1} dx$ を用いて理論曲線が求まる．図 2 のデータを用いて τ_k たちとパラメータ N, a, b を求めることで，

$$(N, a, b) \sim (8 \times 10^5, 5 \times 10^{-4}, 0.8) \quad (31)$$

を得た [22, 24]．これを用いて理論曲線を図 2 の実際のデータに重ねたのが図 3 である．思い切り単

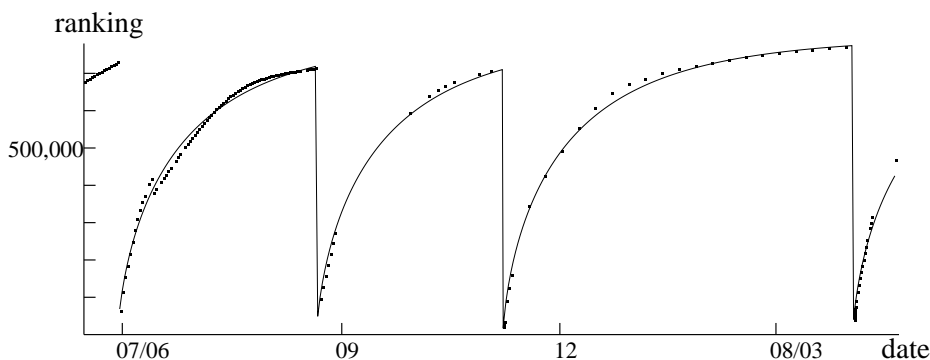


図 3：確率的順位付けモデルの理論曲線の観測値への当てはめ

純な数理モデルにしては，定量的にも現実のランキングのデータをよく説明する，と考える．

ネット書店としての Amazon は，かつてロングテールビジネスの草分け的存在として注目された [1]．ビッグヒットを除く裾野に位置する本は個別にはめったに売れなくても，ほとんどすべての本が裾野に位置するので，合計すれば無視できない売上をもたらすかもしれない，というのがロングテールビジネスの可能性である．ウェブ小売業以前は商品陳列のためのコストが大きかったので，ビッグヒット依存型の商売しかありえなかったのに対して，ウェブ小売業では商品ごとの陳列コストがきわめて小さくロングテールビジネスの可能性が現実的な検討課題になった．商品の売上分布は社外秘に属するが，ランキングという公開情報だけを用いることで Amazon.co.jp の売上分布の近似である (29) と (31) を得た．つまり，ロングテールビジネスであるか否かを部外者でありながら分析できる．

一般化パレート分布は格差を論じる状況でよく用いられると書いたが，分布の形 (29) のとおり， b が小さいとき裾野に比べてビッグヒットの寄与が圧倒的であり， b が大きいときはロングテールが無視できない． N が大きいとき，ロングテールの売上への寄与が全売上の中で無視できるかできないかは

b が 1 より大きいかわ小さいかが判定基準となることがわかる．データを当てはめた結果 (31) から $b < 1$ とわかったので, Amazon.co.jp ではロングテールの売上は無視できる．Amazon はロングテールに非ずである [23] .

2.3 強度が時刻に依存する確率的順位付け模型

強度関数 w_i たちが時刻の関数の場合でも, 位置依存性が無ければ, (2) の $\tilde{v}_i^{(N)}$ が i について独立な強度 $w_i(t)$ の (非一様) ポワソン過程なので, w_i たちが定数の場合と同様に独立確率変数列の大数の法則で極限を直接得て, (非一様) ポワソン過程の確率で書ける [19, 24, 23, 40, 41] . W は時刻の 1 変数連続関数の集合 $C^0([0, T])$ と同一視できて, w_i たちが定数のときの $W = \mathbb{R}_+$ よりも実質は複雑だが, 標準の最大値ノルムに基づく距離について可分完備距離空間だから, 確率論としては変わることなく極限についての確率を議論できる．定常分布という言葉は使いがたいが, 最上流を起点とする ($\gamma = (0, t_0)$) 粒子の軌道 $y_C(\gamma, t)$ については強度の分布 $\lambda(dw) = \mu_0(dw \times [0, 1])$ を用いて定数の場合の軌道 (25) で $w \times (t - t_0) \mapsto \int_{t_0}^t w(s) ds$ と置き換えた

$$y_C(\gamma, t) = 1 - \int_{C^0([0, T])} e^{-\int_{t_0}^t w(s) ds} \lambda(dw), \quad t \geq t_0, \quad \gamma = (0, t_0), \quad (32)$$

が成り立つ [19] . (32) と (27) から強度が時刻依存性を持つ場合への (28) の一般化もわかる .

実質的に独立確率変数列の大数の法則に帰着することから自然なことだが, 経験分布の収束だけでなく (第 1.2 節の定理 2 の後半の結論に相当する), 名札付き粒子過程の独立確率過程列への収束 (カオスの伝搬) が成り立ち [40], さらに, 汎関数中心極限定理 (以上の確率過程の収束の速さが中心極限定理で記述できること) の成立も永幡さんによってわかっている [41] .

強度が時刻に依存する確率的順位付け模型の極限定理は数学としては以上のように明快であるが, たとえば (32) と (27) を組み合わせた公式を強度が定数の場合の図 3 や (31) のように実際の現象と比べようとすると, λ が非負実数 \mathbb{R}_+ 上の分布ではなく連続関数 $C^0([0, T])$ 上の分布なので, 決めるべき母数が急増して, そのままでは現実的な統計的当てはめが困難になる．たとえば社会現象では (気象などの自然現象でも) 活動の昼夜差は無視できない．実際, Amazon ランキングも, 竹島佑介さんが書いたプログラムで自動的にデータを取得できるようになってから, 毎時の更新ごとのデータを取ってみると, 図 4 のように毎日定時のデータ (図 3) に小さな日周期の時間変化を重ねたものに見える．この複雑な状況への簡単な対処として, 関数 $a: [0, T] \rightarrow \mathbb{R}_+$ を決めて強度を $W = \{w \cdot a \mid w \in \mathbb{R}_+\} \subset C^0[0, T]$ に制限 (すべての粒子 i について同一の時刻依存性を仮定) すると, w が定数の場合と同様に λ を \mathbb{R}_+ 上の分布と同一視できて,

$$y_C(\gamma, t) = 1 - \int_{\mathbb{R}_+} e^{-w(A(t) - A(t_0))} \lambda(dw), \quad A(t) = \int_0^t a(s) ds, \quad t \geq t_0, \quad \gamma = (0, t_0), \quad (33)$$

を得る [19, 24] .

さらに a が周期 T_p を持つ ($a(t + T_p) = a(t)$, $t \geq 0$) 場合は, $\frac{1}{T_p} \int_0^{T_p} a(u) du = 1$ を満たすように w の定数倍を取り直せば, $A_p(t) := A(t) - t = \int_0^t (a(u) - 1) du$ は周期 T_p の周期関数だから, 周期 T_p ごとに観測すれば,

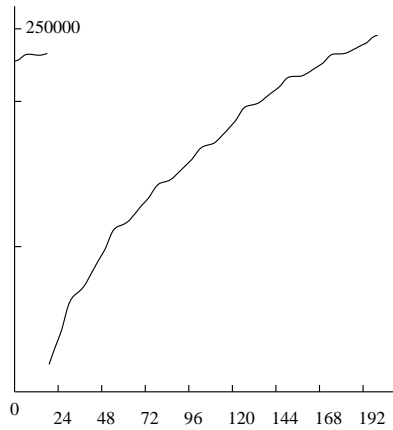


図 4: Amazon ランキングの毎時更新ごとの観測値は、毎日定時の観測値(図 2)に日周期の変化を重ねた形をしている。横軸は時刻 t 、全体で約 8 日、縦軸は X_i 。

$$y_C(\gamma, t_0 + nT_p) = 1 - \int_{\mathbb{R}_+} e^{-w n T_p} \lambda(dw), \quad t \geq t_0, \quad \gamma = (0, t_0), \quad (34)$$

となって、(25) で $t = t_0 + nT_p$ とした場合に一致する。すなわち、日周期がある強度の場合でも、毎日定時のデータは強度が定数の (25) に基づく (28) に当てはめてよい。初期の強度 w が定数の場合の研究 [20, 21] の時点では強度が時刻に依存する場合の証明を持っていなかったし、手動だったので時刻依存性を確認できなかったが、この (34) が成り立つ見通しの下で、毎日定時の観測と強度が定数の場合の (25) に基づいて (31) や図 3 の当てはめ [22] を行っていた。

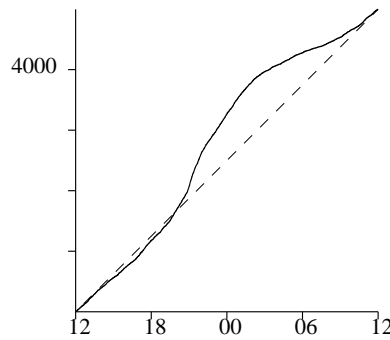


図 5: 累積強度関数 $A(t)$ の例。横軸は正午からの 1 日。強度に時刻依存性がなければ直線(点線)になる。

詳しく [19, 24, 23] は略すが、図 5 が (33) の A の 1 周期 1 日分の例である。強度は A の時間微分なのでこの例では 22 時頃活動が活発だが日が変わると平均未満となり、社会活動の「昼夜逆転」はないことが見られる。

Amazon ランキングに戻って、図 4 のような毎時刻のデータを (33) に当てはめることで、毎日定時のデータを (25) に基づく (28) に当てはめる場合よりも多くのデータが使えるが、(31) と同様に $b = 0.8 < 1$ を得て [24, 23]、Amazon はロングテールに非ずという結論は変わらない。

2.4 強度が位置にも依存する確率的順位付けモデルへ

初期の研究の頃にセミナー等で必ず質問されたことの1つに、ランキングが上位であることの宣伝効果はあるか、ということがあった。この問題に数学的に答えるには強度 $w(x, t)$ が時刻 t だけでなく位置 x にも依存する場合を考える必要がある。現実や応用と切り離しても、時刻だけの関数と同様に証明できるならば1変数で終わるのは不自然であり、難しくなるなどの程度数学が変わるか見る意味がある。実際、難しくなることがわかった [30, 25, 26, 27, 29, 28]。

強度が位置依存性を持つ確率的順位付けモデルの流体力学極限を初めて証明したのは楠岡誠一郎さんとの共著 [30] である。方程式 (1) 自体そのとき初めて楠岡さんが書いて下さったので、(時刻方向の) 劣マルチンゲール不等式を用いた楠岡さんの証明の構図とともに画期的な結果であるが、[30] の定理の仮定に2点の宿題が残った。

(i) 強度 w (は有界閉区間上の連続関数なので個別に有界だが、その) の分布の台が有界であることを [30] で仮定した。いっぽう、現実のウェブのランキングは (29) のように格差が巨大で、極限で有界でないとするのが良い描像である。現実への応用を抜きにしても、強度 w が位置依存性を持たない場合 [20, 21, 19] は、強度の分布に有界性の仮定は不要であった。位置依存性が僅かでも入ったとたんに有界性が必要になる理由は見当たらない。

(ii) 分布の空間は [30] では全変動ノルムに基づく位相を入れた。これは、大数の法則で打ち消し合う揺らぎを、強度 w_i が共通 ($w_i = w$) の粒子の間の揺らぎに限ったことを意味する。(関連するわかりやすい事実として、粒子系のような離散分布が、パレート分布のような連続分布に収束しない。) 他方、位置依存性を持たない場合の証明は、弱収束位相による収束で、異なる w_i たちの間の揺らぎの相殺も許されている。極限としてパレート分布などの連続分布を選ぶことも差し支えない。◇

第3節で証明 [25, 27, 28] の概要を紹介するが、第1.2節の定理2では、強度の分布の台はパレート分布を含む非有界を許し、分布の空間の位相は弱収束位相と同様に異なる強度のポワソン過程の間の揺らぎの相殺を排除せず、離散分布が連続分布に収束することも許す。

3 確率的順位付けモデルの流体力学極限

本節は第1.2節の定理2の [25, 27, 28] に基づく証明の概要である。(第2節を先に読まれた場合は、第1.2節の諸量の定義を以下適宜参照いただきたい。)

3.1 証明の構図

証明の扇の要の位置に、「流れが定める強度に従う確率的順位付けモデル」の発見がある。この名称の前半は、(11) で定義した流れ $\theta \in \Theta_T$ ごとに決まる粒子系であることを意味する。また、この粒子系は (14) で用意した、直前の到着時刻に強度が依存する点過程に基づいて、(3) の類推で定義される。この粒子系が以下の性質を持つことを証明することで、元の粒子系(強度が位置に依存する確率的順位付けモデル)の流体力学極限の存在を証明した。

(i) この系において粒子の先頭への跳びの事象の定義関数(確率変数)とその期待値、すなわち事象の確率、との差が極限で消えること(第3.4節, [27])。

(ii) 上記事象の確率(期待値)も収束すること、したがって、この粒子系の結合経験分布および特性曲線、そして名札付けした粒子への他の粒子の影響が順次収束すること(第3.3節, 第3.2節, [28])。

(iii) 以上の収束は任意の流れ $\theta \in \Theta_T$ に対して成立するが、特に、結合経験分布の極限の特性曲線に等しくなる θ がただ1つあること(第3.5節, [25])。

(iv) 上記の、唯一存在する流れ θ が定める強度に従う確率的順位付けモデルの定義の中の、直前の到着時刻に強度が依存する点過程を、元の粒子系を定義するポワソン乱測度を用いて定義 (coupling) することで、元の粒子系との差が $N \rightarrow \infty$ で消えること (第 3.2 節, 第 3.6 節, [28]). \diamond

流体力学極限における中間模型は局所平衡の概念が周知だが、確率的順位付けモデルは拡散過程でない (非対称過程なので) 平衡系はなく、局所平衡系と流れが定める強度に従う確率的順位付けモデルの間には数学的關係はない. 構図の類似は比喻にとどまるが、流体力学極限と呼ばせていただく. 文献との対応に関して、ポワソン過程の非独立増分な一般化を扱う第 3.2 節の内容と 1 次元流体の偏微分方程式系の非局所項を含む一般化になる第 3.5 節の内容、および、非減少関数値独立確率変数列の大数の一様完全法則を扱う第 3.4 節は確率的順位付けモデルから切り離して閉じた形に結果を纏めたので、それぞれ [25] および [27] に切り離し、その他の部分は [28] に書いた. 切り離した部分については [26, 29] に追加事項がある.

以下、それぞれについて紹介する.

3.2 直前の到着時刻に強度が依存する点過程

直前の到着時刻に強度が依存する点過程 $\tilde{\nu}(t)$ は、主定理証明のための中間模型 (流れが定める強度に従う確率的順位付けモデル) の定義で用いる. 自然に主定理において極限を記述する素材であり、そのため第 1.2 節の (14) で既に紹介した. 名称どおり、2 時刻の関数 $w(s, t)$ を強度関数として、時刻 t での強度が t までの最新の到着時刻 s にも依存する. 強度関数 w が s について定数ならば通常の非一様ポワソン過程である. 応用上は、照明が性能に応じた強度関数に従ってランダムに切れるが、切れた時点で買い換えると、技術の進捗や環境上の規制によって、その時点で買える照明の規格が違うことを考慮した累積交換回数を表す.

ポワソン過程と同様にポワソン乱測度 $\nu(d\xi \times ds)$ を用いて構成できて、種々の公式を得るとともに、粒子系では元の模型と中間模型の coupling を得て、両者の極限の一致も得る [25, 26, 28]. 最後の点について先に立ち入ると、時刻 t までの最後の到着時刻を $\tau^*(t)$ とおくと、

$$\tau^*(t) = \int_{s \in (0, t]} \int_{\xi \in [0, \infty)} (s - \tau^*(s-)) \mathbf{1}_{\xi \in [0, w(\tau^*(s-), s-))} \nu(d\xi \times ds)$$

となる [26, §1.2] ことに注意すると

$$\tilde{\nu}(t) = \int_{s \in (0, t]} \int_{\xi \in [0, \infty)} \mathbf{1}_{\xi \in [0, w(\tau^*(s-), s-))} \nu(d\xi \times ds)$$

を得るので、 ν を元の確率的順位付けモデル (1) のポワソン乱測度に取りれば coupling を得る.

ポワソン過程とは対照的に、独立増分ではなく、公式は複雑になる. たとえば、時間 $(s, t]$ に到着が無い確率は、一時的に $W(u, v) = \int_u^v w(u, s) ds$ とおくと、 $k = 1, 2, 3, \dots$ について、 $u_0 = 0$ として

$$P[\tilde{\nu}(t) = \tilde{\nu}(s) = k] = \int_{0 \leq u_1 \leq u_2 \leq \dots \leq u_k \leq s} e^{-W(u_k, t)} \prod_{i=1}^k w(u_{i-1}, u_i) e^{-W(u_{i-1}, u_i)} du_i$$

となる [25, §3]. ここで、 w が第 1 変数について定数ならば、よく知られた公式

$$\int_{0 \leq u_1 \leq u_2 \leq \dots \leq u_k \leq s} \prod_{i=1}^k f(u_i) du_1 du_2 \dots du_k = \frac{1}{k!} \left(\int_0^s f(v) dv \right)^k, \quad s \geq 0, \quad k = 1, 2, \dots,$$

よって、 $P[\tilde{\nu}(t) = \tilde{\nu}(s)] = e^{-W(s,t)}$ となつて、ポワソン過程の結果を特別な場合として得る。

中間模型の粒子の先頭への跳びというランダムな事象の偏差が消えることは後述の大数の法則の範疇だが、偏差の基準となる期待値（確率）の収束の証明 [28, §3.3.3] と、中間模型を定める流れ $\theta \in \Theta_T$ を決定する関数方程式 (17) の右辺の写像 $\mathcal{G} : \Theta_T \rightarrow \Theta_T$ が縮小写像であることの証明 [25] に上記を含む精密な結果を用いた。第 2.4 節に即して書くと、[30] の強度分布の台が有界という仮定を外すのに以上の精密な評価も用いたということである。なお、もう一方の宿題の収束位相の選択は大数の法則の問題なので、模型固有の詳細評価では解決しない。

3.3 流れが定める強度に従う確率的順位付け模型

流れの集合 (11) から $\theta \in \Theta_T$ を選んで固定し、各粒子番号 i ごとに (13) の $\tilde{w}_{\theta, w_i, y_i^{(N)}}(s, t)$ を (14) の強度とする、直前の到着時刻に強度が依存する点過程を $\tilde{\nu}_i^{(N, \theta)}$ とおく。第 3.2 節に従って元の模型 (1) と $\nu_i^{(N)}$ を通して couple してあるとする（特に、 $\tilde{\nu}_i^{(N, \theta)}$, $i = 1, \dots, N$, は独立な過程である）。元方程式（の書き換え）(3) の Y_i と $\tilde{\nu}_j$ の上の添字を $N \mapsto (N, \theta)$ とした、

$$Y_i^{(N, \theta)}(t) = y_i^{(N)} + \frac{1}{N} \sum_{j=1}^N \int_0^t \mathbf{1}_{Y_j^{(N, \theta)}(s-) > Y_i^{(N, \theta)}(s-)} \tilde{\nu}_j^{(N, \theta)}(ds) - \int_0^t Y_i^{(N, \theta)}(s-) \tilde{\nu}_i^{(N, \theta)}(ds)$$

で定まる $Y_i^{(N, \theta)}$, $i = 1, \dots, N$, を、流れ θ が定める強度に従う確率的順位付け模型と呼ぶ。本来のモデルと同様に、位置強度結合経験分布 $\mu_t^{(N, \theta)}$, 離散特性曲線 $Y_C^{(N, \theta)}$, 分布関数 $\varphi^{(N, \theta)}$ を、それぞれ (4), (5), (6) で $Y_i, \tilde{\nu}_j, Y_C$ の上の添字を $N \mapsto (N, \theta)$ とした式で順次定義すると、（本来の、強度が位置に依存する確率的順位付け模型 (3) と違って、） $\tilde{\nu}_i^{(N, \theta)}$ たちが独立なので、（第 2 節で言及した、強度が位置依存性を持たない確率的順位付け模型と同様に、） $\varphi^{(N, \theta)}$ については先頭に跳ぶ規則に基づく従属性（元の方程式 (1) では右辺第 2 項）を避けることができ、独立確率変数列の大数の法則が使える。位置依存性を持たない確率的順位付け模型との相違は、 $\tilde{\nu}_i^{(N, \theta)}$ は直前の到着時刻が分布することから強度がサンプルごとに変わるので、ポワソン過程ではなくなって確率が複雑になる点である。

もう少し具体的に書くと、 $J_i^{(N, \theta)}(t_0, t) = \{\omega \in \Omega \mid \tilde{\nu}_i^{(N, \theta)}(t)(\omega) > \tilde{\nu}_i^{(N, \theta)}(t_0)(\omega)\}$ と置くと、 $\gamma = (y_0, t_0) \in \Gamma_t$ かつ $t \in [0, T]$ （境界点）に対して、

$$\varphi^{(N, \theta)}(dw, \gamma, t) = \frac{1}{N} \sum_{j; y_j^{(N)} \geq y_0} \mathbf{1}_{J_j^{(N, \theta)}(t_0, t)^c} \delta_{w_j}(dw)$$

となつて、 $\varphi^{(N, \theta)}(dw, (y_0, t_0), t)$ は独立確率変数の算術平均である。この収束を言いたい。サンプル $\omega \in \Omega$ に関しては概収束、すなわち ω を固定することに証明するが、その際、 W 上の分布の収束については弱収束型の位相、すなわち、しかるべき W 上の実連続関数の族からとった h を積分した量

$$\varphi^{(N, \theta)}(h, \gamma, t) = \frac{1}{N} \sum_{j; y_j^{(N)} \geq y_0} \mathbf{1}_{J_j^{(N, \theta)}(t_0, t)^c} h(w_j)$$

で考える．収束の証明は，大数の法則，すなわち，期待値（確率）からの偏差の算術平均の消失

$$\begin{aligned} & \lim_{N \rightarrow \infty} (\varphi^{(N,\theta)}(h, \gamma, t) - E[\varphi^{(N,\theta)}(h, \gamma, t)]) \\ &= \frac{1}{N} \sum_{j; y_j^{(N)} \geq y_0} (\mathbf{1}_{J_j^{(N,\theta)}(t_0, t)^c} - P[J_j^{(N,\theta)}(t_0, t)^c]) h(w_j) = 0 \end{aligned} \quad (35)$$

と，期待値（確率） $\frac{1}{N} \sum_{j; y_j^{(N)} \geq y_0} P[J_j^{(N,\theta)}(t_0, t)^c] h(w_j)$ の収束に，自然に分けられる．後者は証明

全体の中では一番初等的だが，第 3.2 節で言及したとおりモデルの極限の詳細性質を用いる．前者は，確率的な揺らぎの相殺が証明全体の中でここだけにあるので，収束の原動力だが，系の詳細によらないので一般的な枠組の中で次に紹介する．

3.4 関数値独立確率変数列の大数の一様完全法則

中間モデルは本来のモデルとの差の N についての概収束極限を調べることになるが，従属性を評価する際に Gronwall 不等式型の議論に持ち込むために (35) は 2 時刻 t_0 と t についての一様な収束を目標とし，その議論で N の次数を (+0 だが) 損することを見越して (35) では次数を稼ぎたい．始終 2 時刻の空間 $\Delta = \{(t_1, t_2) \in \mathbb{R}^2 \mid 0 \leq t_1 \leq t_2 \leq T\}$ の上の関数 $z: \Delta \rightarrow \mathbb{R}_+$ であって，第 1 変数について減少，第 2 変数について増加，両時刻が等しいとき 0 の関数の集合を \mathcal{D} とおく． \mathcal{D} は，非減少関数の始終時刻間の差と同様の単調性を持つが差で書けないものを含む．

\mathcal{D} に値を取る確率変数列 $Z_i^{(N)}: \Omega \rightarrow \mathcal{D}, i = 1, 2, \dots, N$ ，を考えたい．(35) はその典型例だが，大数の法則として一般化するには少し回りくどい表現が必要で，以下 \mathcal{D} 値独立確率変数列というとき，各時刻ごとに $Z_i^{(N)}(s, t): \Omega \rightarrow \mathbb{R}_+$ が実確率変数で，有理点での実確率変数たちの組が，異なる i の間で独立とする．(一様収束位相で跳びを許す関数を考えるので，通常の実確率変数列の素朴な一般化としてボレル確率変数として扱うと，非可分空間の直積測度の問題が生じる [5]．実確率変数列の大数の法則ではなく Glivenko–Cantelli の定理の一般化と呼ぶのが適切かもしれない [12, 29]．)

定理 3 ([27, Theorem 1]) $r > 0$ および $q > 2$ とする．各 $N \in \mathbb{N}$ に対して $Z_i^{(N)}, i = 1, \dots, N$ ，を \mathcal{D} 値独立確率変数列とし， $M > 0$ と各 i について $w_i^{(N)} \geq 0$ を定数として以下を仮定する．

- (i) $E[Z_i^{(N)}(0, T)^q]^{1/q} \leq M^{(N)}$,
- (ii) $|E[Z_i^{(N)}(t_1, t_2) - Z_i^{(N)}(s_1, s_2)]| \leq M^{(N)} w_i^{(N)} (|t_1 - s_1|^r + |t_2 - s_2|^r)$,
 $(s_1, s_2), (t_1, t_2) \in \Delta$.

このとき算術平均 $Y^{(N)} = \frac{1}{N} \sum_{i=1}^N Z_i^{(N)}$ は， q と r だけで決まる N_0 以上の N に対して

$$E\left[\sup_{(t_1, t_2) \in \Delta} |Y^{(N)}(t_1, t_2) - E[Y^{(N)}(t_1, t_2)]|^q \right] \leq \frac{M^{(N)q} 2^{q-1}}{N^{q^2 r / (2qr + 2r + 2)}} (C_q^q (2T \bar{w}^{(N)})^{1/r} + 1) + 2^{2q}$$

を満たす．ここで， $\bar{w}^{(N)} = \frac{1}{N} \sum_{i=1}^N w_i^{(N)}$ ，そして C_q は q のみで定まる正定数である．さらに $(q^2 - 2q - 2)r > 2$ かつ $\{M^{(N)}, \bar{w}^{(N)}\}$ が有界ならば任意の $\epsilon > 0$ に対して

$$\lim_{N_0 \rightarrow \infty} \sum_{N \geq N_0} P\left[\sup_{(t_1, t_2) \in \Delta} \left| \frac{1}{N} \sum_{i=1}^N (Z_i^{(N)}(t_1, t_2) - E[Z_i^{(N)}(t_1, t_2)]) \right| > \epsilon \right] = 0$$

が成り立つ．

◇

定理 3 の結論の後者の式は、前者のモーメント評価の N についての級数が収束することからチェビシェフの不等式によって得られるが、概収束、すなわち確率 1 の各サンプルに対して算術平均の偏差が 0 に収束する、という主張よりも強い。実確率変数列の収束において [33] が完全収束と名付けて独立実確率変数列の大数の強法則に適用した [33, 13, 14] ので、大数の完全法則と呼ぶことにする。この収束はランダムウォークの概収束と違って、異なる N に対する増分を（同分布な範囲で）取り替えても成り立つ（ i が同じで N の異なる $Z_i^{(N)}$ が同じ確率変数でなくてよい）。ランダムウォークではサンプルごとに、現在位置に 1 歩を追加する「酔歩」なのと対照的である。現実の流体力学極限は有限の N を極限で近似（理想化）するので、完全収束が適切である。

本稿で扱うのは有界な関数列なので q を任意に大きく取れることから、定理 3 のモーメント評価は、評価的に書くと、任意に小さい $\epsilon > 0$ に対して $Y^{(N)} = O(N^{1/2+\epsilon})$ を与える。ここと、第 3.6 節で言及する中間模型と本来の模型の比較の、2 箇所で $O(N^{1/2})$ に届かない。目標の流体力学極限の証明には余裕で十分だが、汎関数中心極限定理は不明である。2 箇所と書いたが、2 時刻について一様な評価を考えるのは第 3.6 節の本来の模型との比較のためなので、自然なことだが、実質は従属性を外すところ 1 箇所に評価の難しさが集中する。定理 3 の N の次数は期待値の t についてのヘルダー連続性の仮定 (ii) の次数 r が関与する。Glivenko–Cantelli の定理の 1 変数関数値確率変数列への直接的な拡張では、ヘルダー連続性の仮定が不要なのと対照的である [27, 29]。

3.5 極限分布を記述する 1 次元多成分流体の偏微分方程式系

第 3.4 節の大数の完全法則と第 3.3 節で言及した期待値の収束の合わせ技が中間模型の収束を担う。後者は第 3.2 節で紹介したような詳細評価を要したが、いずれも流れ $\theta \in \Theta_T$ によらず成立するので、元の模型に収束する中間模型がどの θ かはここまででは決まらない。

第 1.2 節で言及したとおり、本来の確率的順位付け模型と流れが定める強度に従う確率的順位付け模型の 2 つの粒子系は、後者の流れを (17) の \mathcal{G} の唯一の固定点 y_C に選ぶとき、Gronwall 不等式型の議論によって、 $N \rightarrow \infty$ の極限で差が消える。

この、 \mathcal{G} の固定点ということ、あるいは、方程式 (17) の解ということ、を、中間模型の特性曲線の極限（極限の流れ）の言葉で記述すると、「蒸発だけで 1 方向に流れる 1 次元多成分流体を表す偏微分方程式系の蒸発率 w が位置に依存する場合の非局所項（積分項）を持つ一般化」の解を第 3.2 節の確率過程の確率を用いてあらわに書ける、と理解できる。

このことを説明すべく、古典的な偏微分方程式との関係を示すために、強度（流体では蒸発率） w の分布が $\lambda = \sum_{\beta} r_{\beta} \delta_{w_{\beta}}$ の場合に、流体の成分 α の密度 $\frac{d\mu_t(\{w_{\alpha}\} \times dy)}{dt}$ の下流側の積分 $U_{\alpha}(y, t) = \mu_t(\{w_{\alpha}\} \times [y, 1])$ が満たす方程式を書く。まず、 w が位置に依存しない場合は、

$$\frac{\partial U_{\alpha}}{\partial t}(y, t) + \sum_{\beta} w_{\beta}(t) U_{\beta}(y, t) \frac{\partial U_{\alpha}}{\partial y}(y, t) = -w_{\alpha}(t) U_{\alpha}(y, t), \quad (y, t) \in [0, 1] \times [0, T]. \quad (36)$$

となる。たとえば、 w が定数の場合の (26) は（微分 t'_C の表式を用いると）(36) を満たすことを確か

められる。(36) が特性曲線 $\frac{dy_C}{dt}(t) = \sum_{\beta} w_{\beta}(t) U_{\beta}(y_C(\gamma, t))$ を用いて解けることは古典 [8] だが, w が位置 y にも依存する場合への偏微分方程式系 (36) と特性曲線の自然な一般化は, 非局所項を持つ

$$\frac{\partial U_{\alpha}}{\partial t}(y, t) - \sum_{\beta} \int_y^1 w_{\beta}(z, t) \frac{\partial U_{\beta}}{\partial z}(z, t) dz \frac{\partial U_{\alpha}}{\partial y}(y, t) = \int_y^1 w_{\alpha}(z, t) \frac{\partial U_{\alpha}}{\partial z}(z, t) dz, \quad (37)$$

および, $\frac{dy_C}{dt}(t) = - \sum_{\beta} \int_{y_C(t)}^1 w_{\beta}(z, t) \frac{\partial U_{\beta}}{\partial z}(z, t) dz$, であることがわかる [25, §1].

(6) などから推測できるとおり (15) の量は $\varphi_{y_C}(\{w_{\alpha}\}, \gamma, t) = U_{\alpha}(y_C(\gamma, t), t)$ であるが, (15) によって, 直前の到着時刻に強度が依存する点過程で U_{α} を表せて, これが (37) を満たす. 古典的な偏微分方程式の自然な一般化が非局所的であり, その解が古典的なポワソン過程の非独立増分な一般化で表示されるということである [25, 26].

3.6 階層多時刻 Gronwall 型不等式

Gronwall の不等式の原型は周知のとおりである.

命題 4 $T > 0, a \geq 0, c \geq 0$ を実定数とする. $x: [0, T] \rightarrow \mathbb{R}$ が可積分関数で $t \in [0, T]$ に対して $x(t) \leq a + c \int_0^t x(s) ds$ を満たすならば, $x(t) \leq a e^{ct}$, $t \in [0, T]$, が成り立つ. \diamond

これを多時刻に一般化すると次を得る.

命題 5 $T > 0, a \geq 0, c \geq 0$ を実定数, q を正整数とする. $x: [0, T]^q \rightarrow \mathbb{R}$ が可積分関数で $(t_1, \dots, t_q) \in [0, T]^q$ に対して

$$x(t_1, \dots, t_q) \leq a e^{c(t_1 + \dots + t_q)} \frac{1}{q} \sum_{i=1}^q e^{-ct_i} + \frac{c}{q} \sum_{i=1}^q \int_0^{t_i} (x(t_1, \dots, t_q)|_{t_i=u}) du$$

を満たすならば, $x(t_1, \dots, t_q) \leq a e^{c(t_1 + \dots + t_q)}$, $(t_1, \dots, t_q) \in [0, T]^q$, が成り立つ. \diamond

これを変数の個数 q を減らす評価で階層化する一般化として次を得た [28, §4].

定理 6 $T > 0$ と $0 \leq d \leq 1$ を実定数, $x_0 = 1$, 各正整数 q に対して, $a_q \geq 0, b_q \geq 0, c_q \geq 0$ を実定数とし, $x_q: [0, T]^q \rightarrow \mathbb{R}_+$ が可積分関数で $(t_1, \dots, t_q) \in [0, T]^q$ に対して

$$x_q(t_1, \dots, t_q) \leq a_q \sum_{i=1}^q x_{q-1}(t_1, \dots, t'_i, \dots, t_q)^d + b_q \sum_{i=1}^q x_{q-1}(t_1, \dots, t'_i, \dots, t_q) + c_q \sum_{i=1}^q \int_0^{t_i} (x_q(t_1, \dots, t_q)|_{t_i=s}) ds$$

を満たすとする. ここで, 斜線付きの変数はその変数を除く意味である. $\tilde{c}_q = \max_{1 \leq k \leq q} kc_k$, $q \in \mathbb{N}$, と

において, q について帰納的に $g_0 = 1$, $g_q = q(a_q g_{q-1}^d + b_q g_{q-1})$, $q \in \mathbb{N}$. とすると,

$$x_q(t_1, \dots, t_q) \leq g_q e^{\tilde{c}_q(t_1 + \dots + t_q)}, \quad (t_1, \dots, t_q) \in [0, T]^q, \quad q \in \mathbb{N},$$

が成り立つ. \diamond

証明の残りの概略は, 元の確率的順位付けモデルと中間模型において, ある時間内に先頭に跳んだ粒子の集合が一致しない事象の確率を各粒子ごとの事象の定義関数の積の期待値で書いて, (従属性があるので期待値の積にはできないが,) Hölder の不等式で N の小さなべきを損しつつ外して定理 6 を

適用することで Gronwall 不等式型の議論に持ち込む。

4 流行度の順位付け

本稿の研究はウェブで見られるランキングの時間変化の継続的な観測から始まった。いわゆるウェブ小売業の爆発的発展という今世紀初頭を特徴づける社会状況の中で、特定企業に限らず至る所で $N \rightarrow \infty$ 極限が良い近似になり得る巨大な商品項目数のリアルタイムの順位付け更新として軌道 $Y_i(t)$ が広く可視化されたことが、本研究の、特に個々の粒子の軌道を追いかける確率過程としての取り扱いに、現代的な意義を与えた。先頭に跳ぶ規則に関するそれまでの研究は [46] 以来、分布、特に定常分布が中心で、個々の粒子の軌道の時間発展を追いかける動機が少なかったかもしれない。

流行度の順位付けという観点から確率的順位付けモデルの軌道を見ると、最後に売れたものを 1 位とすることは、ほとんどの商品、すなわち期間内にせいぜい 1 点しか売れない商品にとっては、唯一の自然な流行度の定義である。すなわち、流行度の順位付けならば必ずロングテール（売れない、下位）側で先頭に跳ぶ規則に近づくという普遍性がある。必然的に商品ごとの流行度の順位は時間的かつ確率的に変動し揺らぐ確率過程である。どんなに人気無く順位が低くても 1 点売れた瞬間に流行度を 1 位とするのは順位の意味をなさない、という疑問に対しては、ビッグヒット商品は順位が下がり始めてもすぐに注文が入って 1 位に跳ぶので、売れない時間の順位低下を観測するのは難しく、時間軸方向の大数の強法則で順位が安定する、というのが答えである。流行度の順位が確率過程であることは常識ではなかったと思うが、ウェブ以前は表彰台やトップ 10 や読書百選のように順位をビッグヒットにのみ付けていたので流行度の順位が安定な値と思込んでいた、と理解する。

確率的順位付けモデルは確率過程としての流行度の順位付けの時間発展の簡単な数理模型であって、新時代に可視化された巨大なランキングの確率的な挙動を説明する。一方、副産物として、一般に社外秘とされる売上の分布、特にロングテールビジネスの成否をランキングという公開情報で分析できる可能性に言及した。ロングテールの価値は直接的な売上利益ではなく、宣伝効果や文化的意義や可能性の芽である云々は（社会的にロングテールの裾野の中のさらに裾野にいる身なので）承知の上で、数学的な関心や定式化が社会的関心を持ちうる結果をも導ければ、ときおり他分野から連絡をいただく嬉しい経験も得る。自然なことだが、主に出版界と中古品関連業界に注目いただいているようである。Amazon.com の創業者氏に伝わっているという話をかつてセミナーの折に伺ったこともあるが、この社外秘の真偽を確かめる術は持っていない。

強度が位置依存性を持たない確率的順位付けモデルは数学的に相対的にやさしく、実測値との肯定的な対応も得た。順位の宣伝効果のような、実測のいまだない効果の存否を知るべく、強度が位置依存性を持つ場合に一般化すると、難しくなる。現実の社会現象は（ランキングのアルゴリズムを提供会社が恣意的に決められる、というちゃぶ台返しに始まって）繊細な効果を見るのに必ずしも適さないが、現象がまだ見出されていないうちに、数学を先に用意しておくことで、どこか別の場所で現象が見つかったときに議論の混乱を防げれば社会的意義があるかもしれない。

強度分布の非有界性と位置依存性を含めるために中間模型を経由したことは、非独立増分過程という形で従属性が無限粒子極限に残る粒子系としては自然だし、極限の詳細を発見したこともめでたい。ただ、従属確率過程とはいえ大数の法則の証明に極限流体系の詳細な性質を要することは気にする向きもあるかもしれない。対称（拡散）過程や独立増分過程の周辺に比べると、そこから少し外れた問

題に使える精緻かつ一般性のある結果の蓄積が急速に減るか否かは筆者の知識の及ばないところなので、確率的順位付けモデルの中に強いて原因を探すと、強度関数の位置変数についての微分の条件 (16) がある [26]. 巨大な社会格差をモデル化する強度の非有界性と比べて、動機の薄い微分は有界性を仮定したが、(16) は第 3.5 節で紹介した意味で極限の流れを定める (17) の \mathcal{G} が縮小写像になることを保証し、極限流 (定理 1 の解) の一意存在に至る. 試みにこの (16) を振動の有界性

$$C'_W := \sup_{w \in W} \sup_{(y,t), (y',t') \in [0,1] \times [0,T]} |w(y,t) - w(y',t')| < \infty,$$

に置き換えると \mathcal{G} の縮小写像性は言えないが、位相的固定点定理は適用できて解の存在は言える [25, Appendix]. ある流体成分の強度がある時空間の小さな領域だけ極端に大きい場合 (塩田の樋の一部を短時間強く加熱する思考実験), その成分がその時空間に届くと急速な蒸発で上流側の流れが激しくなる一方、初期値の僅かな違いでその時空間に届かず流れが遅い状況を考えてとき、対応する粒子系の収束は複雑な問題かもしれない. ただし筆者はそれ以上の結論を持ち合わせていない.

謝辞. 共著者の方々, 掲載の判断をくださった雑誌の編集長と査読者の方々, 著書および本記事を含む記事執筆を支持してくださった出版社および編集委員会等の関係各位と読者の皆様, 研究会やセミナー等の発表機会をくださった方々と聞いてくださった皆様, 分野や大学の枠を越えて私の研究を見つけてくださって勉強し, 何かの折に引用してくださった方々や問い合わせはるばる訪問して下さるなどの形でお話をする機会をくださった方々, 科研費採択の形でチャンスを与え続けてくださった審査員各位, 多くの方々に本研究は支えられてきました. 心から感謝申し上げます. 本研究は現在科研費基盤研究 (C)18K03344 の支援を受けています.

文 献

- [1] C. Anderson, *The Long Tail: Why the Future of Business Is Selling Less of More*, Hyperion Books, 2006.
- [2] J. Barrera, J. Fontbona, *The limiting move-to-front search-cost in law of large numbers asymptotic regimes*, *Annals of Applied Probability* **20-2** (2010) 722–752.
- [3] H. Bauer, *Measure and integration theory*, De Gruyter Studies in Mathematics **26**, Walter de Gruyter & Co., Berlin, 2001.
- [4] M. S. Berger, *Non-linearity and functional analysis*, Academic Press, 1977.
- [5] P. Billingsley, *Convergence of probability measures*, 2nd ed., John Wiley and Sons, New York, 1999.
- [6] J. R. Bitner, *Heuristics that dynamically organize data structures*, *SIAM Journal on Computing* **8** (1979) 82–110.
- [7] G. Blom and L. Holst, *Embedding procedures for discrete problems in probability*, *The Mathematical Scientist* **16** (1991) 29–40.
- [8] A. Bressan, *Hyperbolic systems of conservation laws*, *The one-dimensional Cauchy problem*, Oxford Univ. Press, Oxford, 2005.
- [9] P. J. Burville and J. F. C. Kingman, *On a model for storage and search*, *Journal of Applied Probability* **10** (1973) 697–701.
- [10] Y. S. Chow, H. Teicher, *Probability theory, independence, interchangeability, martingales*, 3rd ed., Springer, 2003.
- [11] F. R. K. Chung, D. J. Hajela and P. D. Seymour, *Self-organizing sequential search and Hilbert's inequalities*, *Journal of Computer and System Sciences* **36** (1988) 148–157.
- [12] R. M. Dudley, *Uniform Central Limit Theorems*, 2nd ed., 2014, Cambridge UP, New York, Cambridge Studies in Advanced Mathematics 0950-6330.
- [13] P. Erdős, *On a Theorem of Hsu and Robbins*, *Annals of Mathematical Statistics* **20** (1949) 286–291.
- [14] P. Erdős, *Remark on my Paper 'On a Theorem of Hsu and Robbins'*, *Annals of Mathematical Statistics* **21** (1950) 138.
- [15] R. Fagin, *Asymptotic miss ratios over independent references*, *Journal of Computer and System Sciences* **14** (1977) 222–250.
- [16] J. A. Fill, *An exact formula for the move-to-front rule for self-organizing lists*, *Journal of Theoretical Probability* **9** (1996) 113–160.
- [17] J. A. Fill, *Limits and rates of convergence for the distribution of search cost under the move-*

- to-front rule*, Theoretical Computer Science **164** (1996) 185–206.
- [18] J. A. Fill and L. Holst, *On the distribution of search cost for the move-to-front rule*, Random Structures and Algorithms **8** (1996) 179–186.
- [19] Y. Hariya, K. Hattori, T. Hattori, Y. Nagahata, Y. Takeshima, T. Kobayashi, *Stochastic ranking process with time dependent intensities*, Tohoku Mathematical Journal **63** (1) (2011) 77–111.
- [20] K. Hattori and T. Hattori, *Existence of an infinite particle limit of stochastic ranking process*, Stochastic Processes and their Applications **119** (2009) 966–979.
- [21] K. Hattori and T. Hattori, *Equation of motion for incompressible mixed fluid driven by evaporation and its application to online rankings*, Funkcialaj Ekvacioj **52** (2009) 301–319.
- [22] K. Hattori and T. Hattori, *Sales ranks, Burgers-like equations, and least-recently-used caching*, RIMS Kokyuroku Bessatsu **B21** (2010) 149–162.
- [23] 服部哲弥, 「Amazon ランキングの謎を解く - 確率的な順位付けが教える売上の構造」, 化学同人出版, 2011.
<http://web.econ.keio.ac.jp/staff/hattori/amazonj.htm>
<http://web.econ.keio.ac.jp/staff/hattori/amazone.htm>
- [24] T. Hattori, *Stochastic ranking process and web ranking numbers*, in *Mathematical Quantum Field Theory and Renormalization Theory*, T. Hara, T. Matsui, F. Hiroshima, eds., Math-for-Industry Lecture Note Series **30** (2011) 178–191.
- [25] T. Hattori, *Point process with last-arrival-time dependent intensity and 1-dimensional incompressible fluid system with evaporation*, Funkcialaj Ekvacioj **60** (2017) 171–212.
- [26] T. Hattori, *Open problems to an infinite system of quasi-linear partial differential equations with non-local terms*, Symposium on Probability Theory 2014/12/16–19, RIMS Kokyuroku **1952** (2015) 9–16.
- [27] T. Hattori, *Doubly uniform complete law of large numbers for independent point processes*, Journal of Mathematical Sciences the University of Tokyo **25** (2018) 171–192.
- [28] T. Hattori, *Cancellation of fluctuation in stochastic ranking process with space-time dependent intensities*, Tohoku Mathematical Journal, to appear, <http://arxiv.org/abs/1612.09398>.
- [29] 服部哲弥, 「確率変数の収束と大数の完全法則 - 少しマニアックな確率論入門」, 共立出版, 2019.
- [30] T. Hattori, S. Kusuoka, *Stochastic ranking process with space-time dependent intensities*, Latin American Journal of Probability and Mathematical Statistics **9** (2) (2012) 571–607. ALEA, Lat. Am. J. Probab. Math. Stat. **9** (2) (2012) 571–607.
- [31] W. J. Hendricks, *The stationary distribution of an interesting Markov chains*, Journal of Applied Probability **9** (1972) 231–233.
- [32] R. Hirade, T. Osogami, *Analysis of page replacement policies in the fluid limit*, Operations Research **58-4** (2010) 971–984.
- [33] P. L. Hsu, H. Robbins, *Complete convergence and the law of large numbers*, Proceedings of the National Academy of Sciences of the United States of America **33** (1947) 25–31.
- [34] N. Ikeda, S. Watanabe, *Stochastic differential equations and diffusion processes*, 2nd ed., North Holland, 1989.
- [35] P. R. Jelenković, *Asymptotic approximation of the move-to-front search cost distribution and least-recently used caching fault probabilities*, Annals of Applied Probability **9** (1999) 430–464.
- [36] P. R. Jelenković, A. Radovanović, *The Persistent-Access-Caching algorithm*, Random Structures and Algorithms, **33-2** (2008) 219–251.
- [37] J. F. C. Kingman, *Random discrete distributions*, Journal of the Royal Statistical Society Ser. B **37** (1975) 1–22.
- [38] G. Letac, *Transience and recurrence of an interesting Markov chain*, Journal of Applied Probability **11** (1974) 818–824.
- [39] J. McCabe, *On serial files with relocatable records*, Operations Research **13** (1965) 609–618.
- [40] Y. Nagahata, *Tagged particle dynamics in stochastic ranking process*, Kodai Mathematical Journal **36** (3) (2013) 397–408.
- [41] Y. Nagahata, *Functional central limit theorem for tagged particle dynamics in stochastic ranking process*, Kodai Mathematical Journal **36** (3) (2013) 409–427.
- [42] R. Rivest, *On self-organizing sequential search heuristics*, Communications of the ACM **19** (1976) 63–67.
- [43] E. R. Rodrigues, *Convergence to stationary state for a Markov move-to-front scheme*, Journal of Applied Probability **32** (1976) 768–776.
- [44] T. Sugimoto, N. Miyoshi, *On the asymptotics of fault probability in least-recently-used caching with Zipf-type request distribution*, Random Structures and Algorithms **29** (2006) 296–323.
- [45] R. L. Taylor, T. C. Hu, *Strong laws of large numbers for arrays of rowwise independent random elements*, International Journal of Mathematics and Mathematical Sciences **10** (1987) 805–814.
- [46] M. L. Tsetlin, *Finite automata and models of simple forms of behaviour*, Russian Mathematical Surveys **18** (1963) 1–27.

(2019年1月4日提出)

確率的順位付けの数理モデル

21

(はっとり てつや・慶應義塾大学経済学部)