

流行度の順位付け

服部哲弥（慶應大・経済）

2012.11.17 OR学会待ち行列研究部会例会（東京工業大学）

START

↑

0 . 目次

- ・先頭に跳ぶ規則の流体力学的極限（位置ジャンプ率結合分布の，粒子数を大きくするスケール極限）の存在定理（数理）
- ・流行度の順位付け（ランキング）のモデルとして，先頭に跳ぶ規則を採用できる（提案）
- ・Amazon.co.jpのランキング，2ch.netのスレッド一覧への当てはめ「Amazonはロングテールに非ず」（応用）

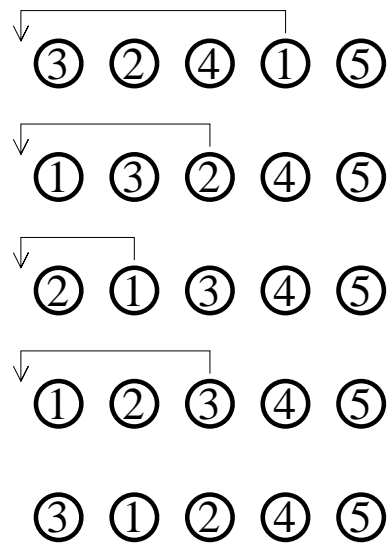
服部哲弥「確率的ランキング - 流行度の順位付けとロングテール分析」日本OR学会誌 57-6 (2012.6) 3-8.

1 . 先頭に跳ぶ規則の流体力学極限

先頭に跳ぶ (Move-To-Front) 規則 : M.L.Tsetlin (1963)

N 自然数

N 個の粒子を一系列に並べた系の並び方についてのランダムな時間発展



どれかの粒子がランダムに先頭に跳び, 追い越された粒子は順位を1ずつ下げる

図は, 粒子 1, 2, 1, 3 の順に跳んだときの並び方の時間発展

積ん読, 超整理法, 最後に跳んだ順, ...

確率順位付け模型

個々の粒子の順位 (ランキング) の時間発展 一般化の方向

ゼッケン番号 i の粒子の時刻 t での位置 (ランキング) $X_i^{(N)}(t)$

$t \geq 0, i = 1, 2, \dots, N, X_i^{(N)} : \Omega \times \mathbb{R}_+ \rightarrow \{1, 2, \dots, N\}$

$X^{(N)} = (X_1^{(N)}, \dots, X_N^{(N)})$

初期値 : $X_i^{(N)}(0) = x_i^{(N)}, i = 1, 2, \dots, N$

時間発展 (確率過程):

(0) **先頭に跳ぶ規則** : ある粒子が先頭 ($X_i^{(N)}(t) = 1$) に跳び, 追い越

された粒子は順位を1ずつ下げる ($X_j^{(N)}(t) = X_j^{(N)}(t-) + 1$)

先頭に跳ぶランダム時刻で決まる:

現実のランキングのモデルとして, できる限り**簡単**なものを選ぶ

ポワッソンのランダムな測度

(1) 先頭に跳ぶ時刻は i について独立

(粒子間で勝手に先頭に跳ぶ：ポワッソンなので，跳ぶ時刻が一致する確率は0)

(2) 時刻 t に位置 x にいる粒子 i は「 $w_i^{(N)}(x, t)$ の頻度」で先頭に跳ぶ

$w_i^{(N)}$: 粒子 i 毎に選んだ2変数非負値関数 (強度, ジャンプ率)

(2-1) 位置依存性がないとき ($w_i^{(N)} = w_i^{(N)}(t)$):

時刻 t までに先頭に跳ぶ回数は強度 $w_i^{(N)}(t)$ のポワッソン過程

すなわち，粒子 i が時間 $(s, t]$ の間に先頭に跳ぶ回数は平均 $\int_s^t w_i^{(N)}(u) du$ のポワッソン分布で，共通部分を持たない時間区間の先頭に跳ぶ回数は独立

(2-2) 位置依存性： \mathbb{R}_+^2 上の一様なポワッソンランダム測度 $\nu_i^{(N)}$ に

$\mathbf{1}_{\xi \leq w_j^{(N)}(X_j^{(N)}(s-), s)}$ をかけて確率積分 ($\mathbf{1}_A$ は事象 A の定義関数)

(古典的な確率積分 (池田 - 渡辺など) の定義に当てはまる)

確率順位付け模型

$$i = 1, 2, \dots, N, t \geq 0$$

$\nu_i^{(N)}$: \mathbb{R}_+^2 上のポワソンランダム測度 ; 強度 $d\xi ds$, i について独立

$$X_i^{(N)}(t) = x_i^{(N)}$$

$$+ \sum_{j=1}^N \int_{s \in (0, t]} \int_{\xi \in \mathbb{R}_+} \mathbf{1}_{X_i^{(N)}(s-) < X_j^{(N)}(s-)} \mathbf{1}_{\xi \leq w_j^{(N)}(X_j^{(N)}(s-), s)} \nu_j^{(N)}(d\xi ds)$$

(下位の粒子の先頭へのジャンプに押される変化)

$$+ \int_{s \in (0, t]} \int_{\xi \in \mathbb{R}_+} (1 - X_i^{(N)}(s-)) \mathbf{1}_{\xi \leq w_i^{(N)}(X_i^{(N)}(s-), s)} \nu_i^{(N)}(d\xi ds)$$

(先頭へのランダムなジャンプ)

ジャンプ率関数の（現時点での）仮定

$$w_i^{(N)}(k, t) = w_i\left(\frac{k-1}{N}, t\right), \quad (\text{極限定理で空間をスケールするので})$$

$$W = \{w_i \in C^1([0, 1] \times \mathbb{R}_+, \mathbb{R}_+) \mid i = 1, 2, \dots\};$$

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

注 . 有界の仮定は $\int \|w\| \lambda(dw) < \infty$ 程度に弱められてもしかるべきだが

大数の法則

- 位置 $Y_i^{(N)} = \frac{1}{N} (X_i^{(N)} - 1)$ とジャンプ率の結合経験分布

$$\mu_t^{(N)} = \frac{1}{N} \sum_{i=1}^N \delta_{(w_i^{(N)}, Y_i^{(N)}(t))} \quad \text{の分布関数 } U^{(N)}(\cdot, y, t) = \mu_t^{(N)}(\cdot, [y, 1))$$

U が良い量

定理 . $N \rightarrow \infty$ で , $U^{(N)}(dw, y, 0) \rightarrow \exists U_0(dw, y)$ (概収束 , y 一様 , W :全変動ノルム) ならば , $U_t^{(N)}(dw, y, t) \rightarrow \exists U(dw, y, t)$ ($dw, y, t \leq T$ 一様) . U は非ランダムな分布 (偏微分 - 積分方程式の解 : 後述) .

さらに , $\frac{1}{N}(x_i^{(N)} - 1) \rightarrow y_i \in [0, 1)$, $i = 1, 2, \dots, L$ ならば ,

$(Y_1^{(N)}(t), \dots, Y_L^{(N)}(t)) \rightarrow \exists (Y_1(t), \dots, Y_L(t))$ (概収束 ; $t \leq T$ 一様)

Y_i は U と ν_i で書ける確率微分方程式の解 (propagation of chaos)



流体力学的極限（極限の特徴づけ）

初期値への仮定： $U_0(dw, y) = \exists \mu_0(dw, [y, 1])$ (y についての分布関数)

$\mu_0(W, dy) = dy$ ($y =$ 相対順位)

$\exists C^0$ 密度 $\exists \sigma = \frac{d\mu_0}{d\lambda \times dy}$; $\mu_0(dw, [0, 1]) = \lambda(dw)$,

$\Gamma_i = \{(y_0, 0)\}$ 初期点, $\Gamma_b = \{(0, t_0)\}$ 境界点,

$\Gamma_t = \{(y_0, t_0) \in \Gamma_i \cup \Gamma_b \mid t_0 \leq t\}$

定理. $y_C : \{(\gamma, t) \mid t \geq 0, \gamma \in \Gamma_t\} \rightarrow [0, 1]$ と

$U(dw, y, t) : [0, 1] \times \mathbb{R}_+ \rightarrow \{W \text{ 上の Borel 測度}\}$,

で以下を満たすものがある:

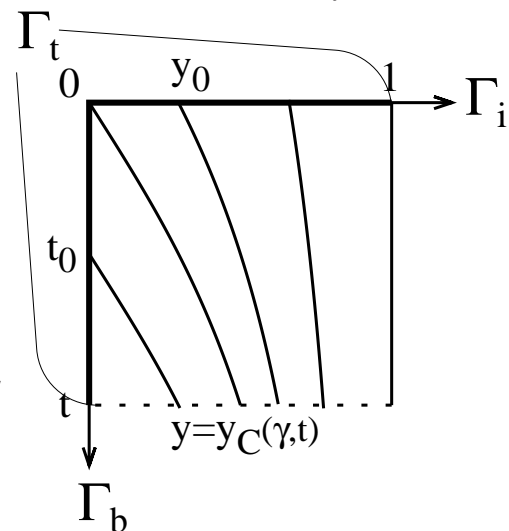
$y_C, \dot{y}_C \in C^0$, $y_C(\Gamma_t, t) = [0, 1]$, $t \geq 0$ (surj.),

$(\forall T)(\forall h : W \rightarrow \mathbb{R} \text{ bdd. conti.}) U(h, y, t) := \int_W h(w) U(dw, y, t)$ Lip. in

$(y, t) \in [0, 1] \times [0, T]$, $U(h, y, t) \uparrow$ in y ($h \geq 0$)

$\gamma = (y_0, t_0) \in \Gamma_t$ に対して, $y_C(\gamma, t_0) = y_0$, $U(dw, y_0, t_0) = U_0(dw, y_0)$

方程式は次ページ



極限結合経験分布を特徴づける微分 - 積分方程式

$$\frac{\partial U}{\partial t}(h, y, t) + V(1_W, y, t) \frac{\partial U}{\partial y}(h, y, t) = -V(h, y, t)$$

$$V(h, y, t) = - \int_y^1 \int_W h(w) w(z, t) \frac{\partial U}{\partial z}(dw, z, t) dz \quad \diamond$$

1階準線形偏微分方程式系 ($U(h, y, t) \in C^1$ の場合) . 本来の定理は積分方程式で書いて, broad (Lipschitz) solution)

V : non-local (ジャンプ率の空間依存性)

空間依存性があると, 本質的に従属変数の大数の法則, 極限も non-local で難

1 階準線形 PDE と特性曲線の方法

ジャンプ率に位置依存性が無い場合

(簡単のため, λ が有限点分布の場合)

$$\frac{\partial U_\alpha}{\partial t}(y, t) + \sum_{\beta} w_\beta U_\beta(y, t) \frac{\partial U_\alpha}{\partial y}(y, t) = -w_\alpha U_\alpha(y, t)$$

境界条件 $U_\alpha(0, t) = r_\alpha$, 初期値 $U_\alpha(\cdot, 0) = u_\alpha(\cdot)$,

位置依存性がないときは積分方程式にはならず, PDE

- 主部が共通の 1 階準線形 PDE 特性曲線 $y = y_C(\gamma, t)$;

位置依存性があると陽には解けないが, 特性曲線の方法を拡張可能

$$\frac{\partial y_C}{\partial t}(\gamma, t) = V(1_W, y_C(\gamma, t), t), \quad t \geq t_0, \quad \gamma = (y_0, t_0) \in \Gamma_i \cup \Gamma_t,$$

特性曲線が満たす積分方程式

命題 . C^1 唯一解 $f : [0, 1] \times \mathbb{R}_+ \rightarrow [0, 1]$ がある :

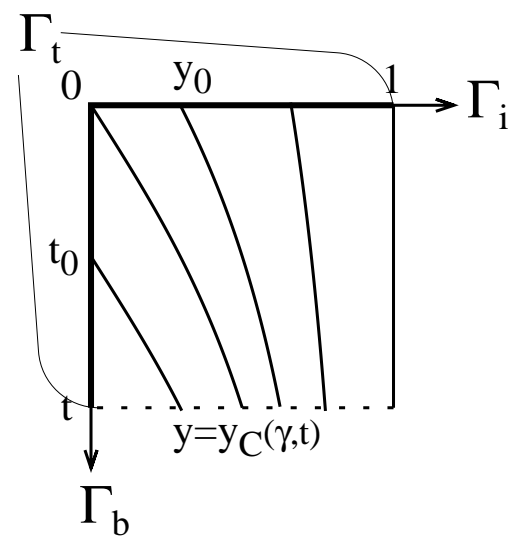
$$f(y, t) = 1 -$$

$$\int_W \left(\int_y^1 \sigma(w, z) \exp\left(-\int_0^t w(f(z, s), s) ds\right) dz \right) \lambda(dw) \quad \diamond$$

$$\sigma = \frac{d\mu_0}{d\lambda \times dy}$$

$\gamma = (y_0, 0) \in \Gamma_i \subset \Gamma$ に対して $y_C(\gamma, t) = f(y_0, t)$

• $(0, t_0) \in \Gamma_b \subset \Gamma$ を始点とする特性曲線はもっと**難しい**



特性曲線と粒子の軌道

$$N < \infty \text{ の対応 : } Y_C^{(N)}(\gamma, t) = y_0 + \frac{1}{N} \sum_{i; Y_i^{(N)}(t_0) \geq y_0} \mathbf{1}_{J_i^{(N)}(t_0, t)}$$

$$J_i^{(N)}(t_0, t) = \left\{ \int_{t_0}^t \int_{\mathbb{R}_+} \mathbf{1}_{\xi < w_i^{(N)}(X_j^{(N)}(s-), s)} \nu_i^{(N)}(d\xi ds) > 0 \right\}$$

粒子 i が時間 $(t_0, t]$ の間に先頭に跳ぶ事象

$Y_C^{(N)}((0, 0), t)$ はジャンプ済み粒子と未ジャンプ粒子の境界

- 定理と同時に $Y_C^{(N)}(\gamma, t) \rightarrow y_C(\gamma, t)$ が証明される
- 位置依存性無い場合と違って, 従属確率変数の大数の法則 **難**

数学的結果の要約

先頭に跳ぶ規則の流体力学的極限の存在定理

- 巨視的な分布（極限分布）の動きは決定論的
- 蒸発項で動く inviscid Burgers 型（積分 - ）偏微分方程式系
- 時刻・位置依存性を入れられる（昼夜差，上位注目効果）
- 特性曲線（1 位に跳ぶ以外の順位変化） 以下で使う

2 . Amazon ランキングの謎を解く

The screenshot shows the Amazon.co.jp product page for the book '統計と確率の基礎 (単行本)' by 服部 哲弥. The page includes a search bar, navigation tabs, a product image, and detailed information.

Amazon.co.jp: 統計と確率の基礎: 服部 哲弥: 本 - Windows Internet Explorer

検索 和書

和書 詳細検索 ジャンル 新刊・予約 ベストセラー ハリー・ポッター 雑誌

統計と確率の基礎 (単行本)
服部 哲弥 (著)
★★★★☆ (2件のカスタマーレビュー)

価格: ¥ 2,100 (税込) この商品は1500円以上国内配送料無料を利用して配送されます。

在庫状況(詳しくはこちら): 在庫あり。この商品は、Amazon.co.jp が販売、発送します。

1点在庫あり。ご注文はお早めに。

出版社: 学術図書出版社: 第2版 (2006/11/10)
ISBN-10: 4873618428
ISBN-13: 978-4873618425
発売日: 2006/11/10
商品の寸法: 21 x 14.8 x 1.6 cm
おすすめ度: ★★★★★ (2件のカスタマーレビュー)
Amazon.co.jp ランキング: 本で159,509位

amazon.co.jp® Amazon.co.jp ホーム

Amazon.co.jp

本のページ中程やや下
Amazon.co.jp ランキング

「Amazon の謎順位。」

‘Internet retailers are extremely hesitant about releasing specific sales data’

謎ならば答えを探そう

ランキングの時間変化

- ・ 本を書くと，自分の本の順位が気になる．

Amazon.co.jp ランキング: 本で373,406位

(1時間後) Amazon.co.jp ランキング: 本で373,977位

(2時間後) Amazon.co.jp ランキング: 本で374,693位

- ・ が，漫然と眺めても1時間に1度更新されることくらいしか分からない

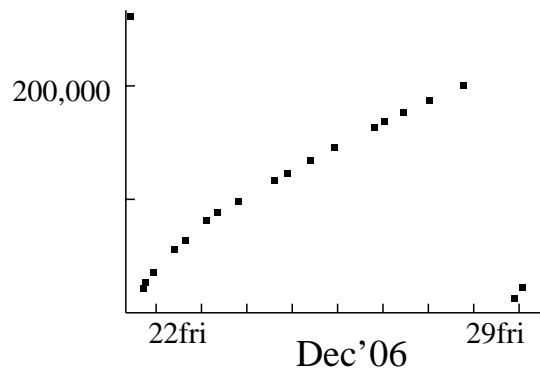
- ・ 方針を**逆転**，

先に簡単な数理モデルを立てて，データの特徴と比較

ランキング = 流行度を反映する順位 理想化：

売れた瞬間を1位にする先頭に跳ぶ規則（確率順位付けモデル）

アマゾン書店での順位の大きな跳び



1回の更新での大きな順位改善の確認

- ・先頭に跳ぶ規則は少なくとも定性的にアマゾン書店ランキングを説明

あり得る疑問：

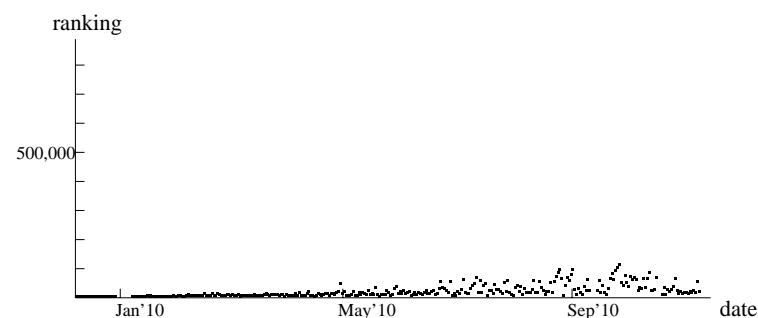
- ・順位はもっと安定的（確定的）な数値であるべきだろう
『売れない専門書が一冊売れただけで一位にな ... るならば，ランキングの意味をなさない』（原啓介・数理科学2011.11）
- ・アマゾン書店がそんな単純なアルゴリズムですますまい
定量的には到底役立たない？
- ・事実：合う（後述）

確率的な順位付けの普遍性

考察：巨大な系の流行度の順位は，大多数にとって確率的（**普遍性**）

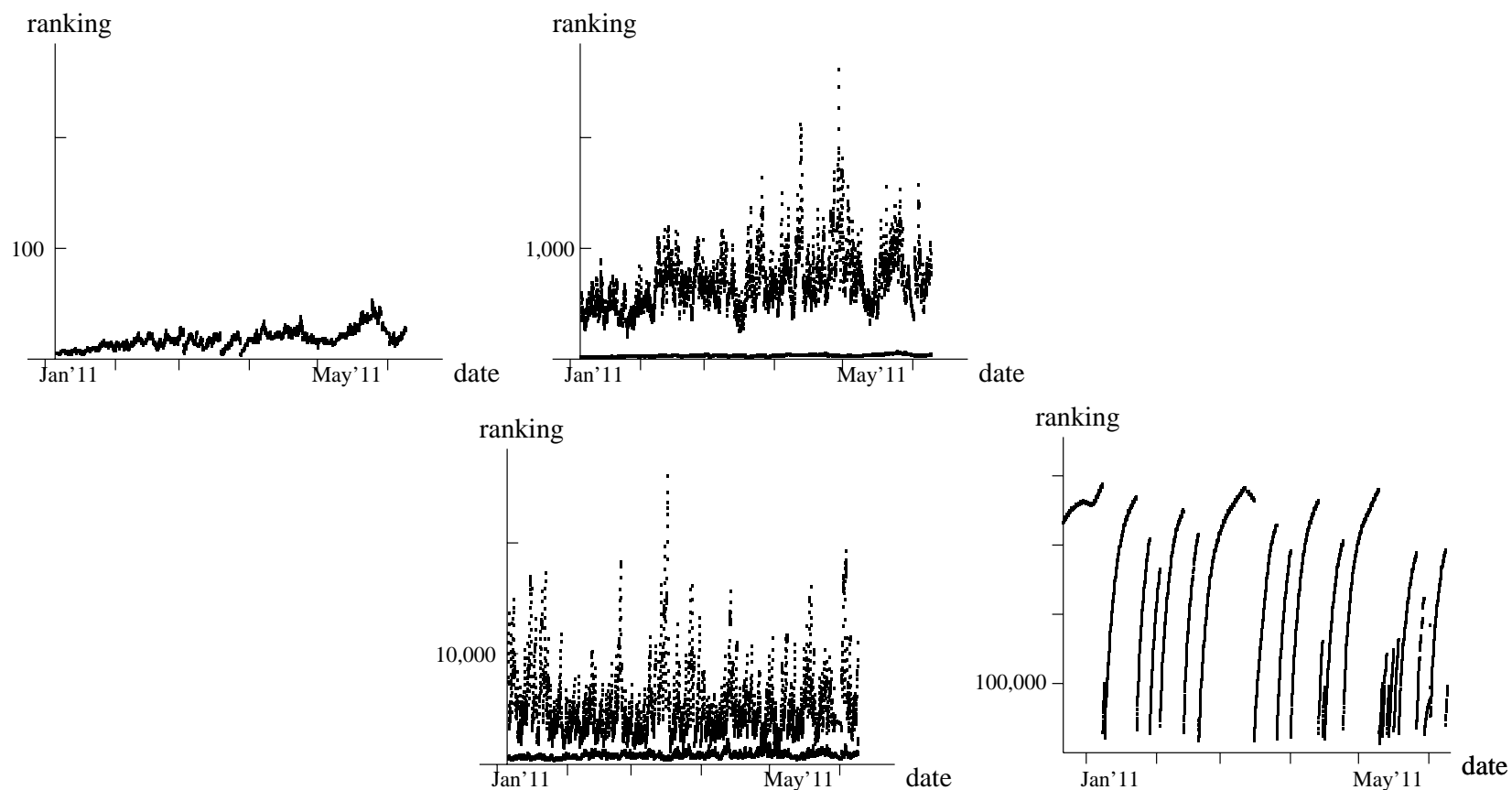
・順位が安定しているのはよく売れる一部の上位だけ．**大多数**の本は数ヶ月に1回以下しか売れない．それらの本の**流行度の判断基準は直近の1回の購入しかない**．

・アマゾン書店のアルゴリズムは，単純な先頭に跳ぶ規則でなく，少なくとも数時間前までの購入状況を反映（澤原昇・MONOQLO 2011.11）．しかし，数ヶ月に1回以下しか売れない**大多数**の本にとっては関係ない．



常識的に順位を考える領域を1位と同一視する縮尺（**新しい現象**）

大数の法則とサンプル



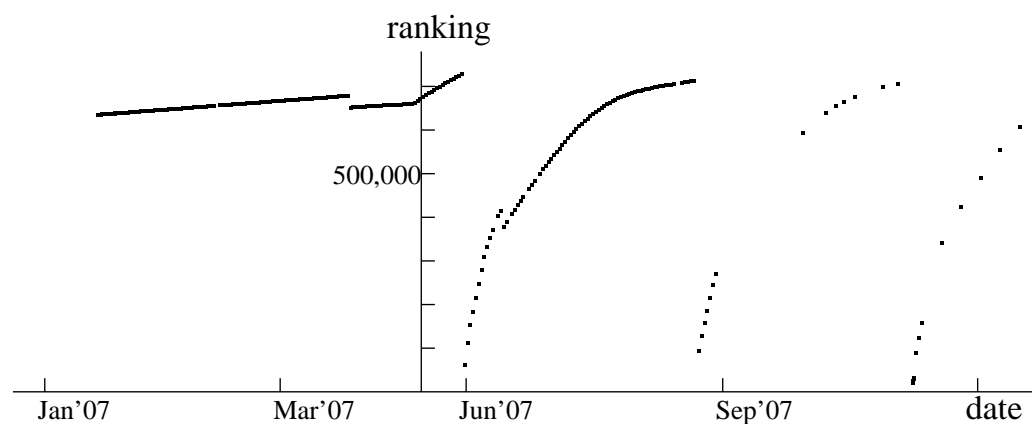
ビッグヒット vs. 普通の本（ロングテール側）
売上についての大数の法則 vs. 確率的順位付け

提案の要約

- ・流行度の順位付け (**ランキング**) のモデルとしての確率順位付けモデル
 - 流行度は確率的
 - 普遍性

 - ・順位付けは確率的な概念である。人々が見てきた『確かな』順位は上位における**多数の支持**という大数の法則の結果である。
 - ・粒子数が大きいと、流行度をどう定義しても (アルゴリズムをどういじっても) **大多数の粒子**の動きはランキングの巨視的振舞いという大数の法則すなわち確率順位付けモデル
- ロングテール**は順位付けが確率的であることが顕著な領域を扱う (**後述**)

3 . 確率順位付けモデルのデータへの当てはめ



(めったに売れない本で) ランキングをしつこく観察記録

軌道 $X_i^{(N)}(t) = X_C^{(N)}(t)$ 特性曲線 $NY_C(t)$ に統計的当てはめ

(初期は) データを人力で... : 毎日21時, 逃したら諦める
昼夜差の除去 (w を時空依存性無しとして当てはめ)

特性曲線（ジャンプ率が定数の場合）

時刻0に1位の本が次に売れるまでのランキング（ w_i が定数の場合）

$$X_i^{(N)}(t) \doteq N Y_C(O, t) = N - N \int_W e^{-w t} \lambda(dw)$$

ランキングの時間変化のデータ $X_i^{(N)}$ （ラプラス逆変換）
売上の頻度の分布（売れる本売れない本の多寡） λ

- ・ **部外秘**に属するだろう情報を，**公開情報**だけで合法的に実証できる！

Pareto分布

ラプラス変換 現実には分布（族）を決めてパラメータ当てはめ
できるだけ単純な族として（一般化）Zipfの法則(Pareto分布)

$$w_i^{(N)} = a \left(\frac{N}{i} \right)^{1/b}; a: \text{最低収入}, b: \text{平等性の指数}$$

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

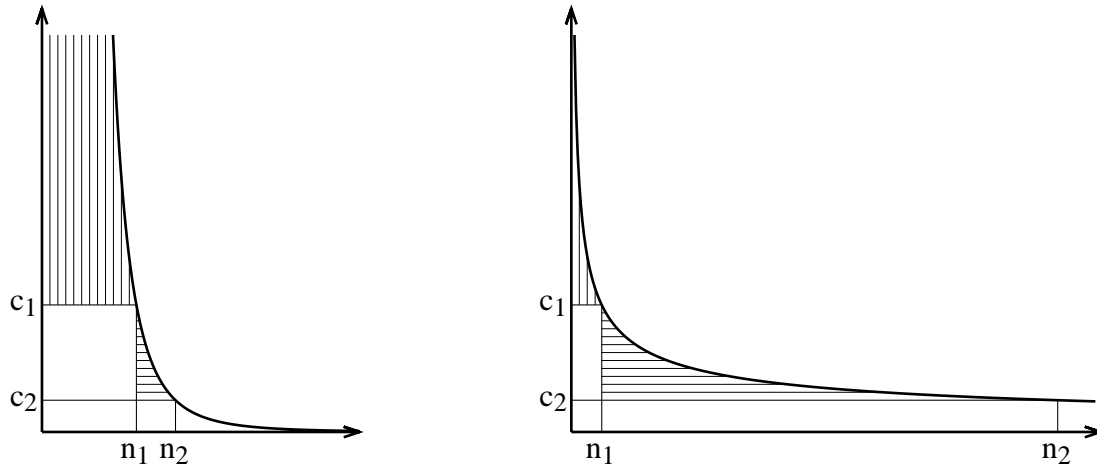
$$\begin{aligned} X_i^{(N)}(t) &\doteq N Y_C(O, t) = N - N \int_W e^{-wt} \lambda(dw) \\ &= N - Nb(at)^b \Gamma(-b, at); \Gamma(z, p) = \int_p^\infty e^{-x} x^{z-1} dx \end{aligned}$$

• N, a, b を与えれば決まる（注． N はAmazon公称ではなく実効値）

平等性の指数 b とロングテール

多品目商品を扱う小売業（例：大型書店） $X_i^{(N)}$ ランキングデータ

i 商品（本のタイトル）, $w_i^{(N)}$ 単位時間当売上,

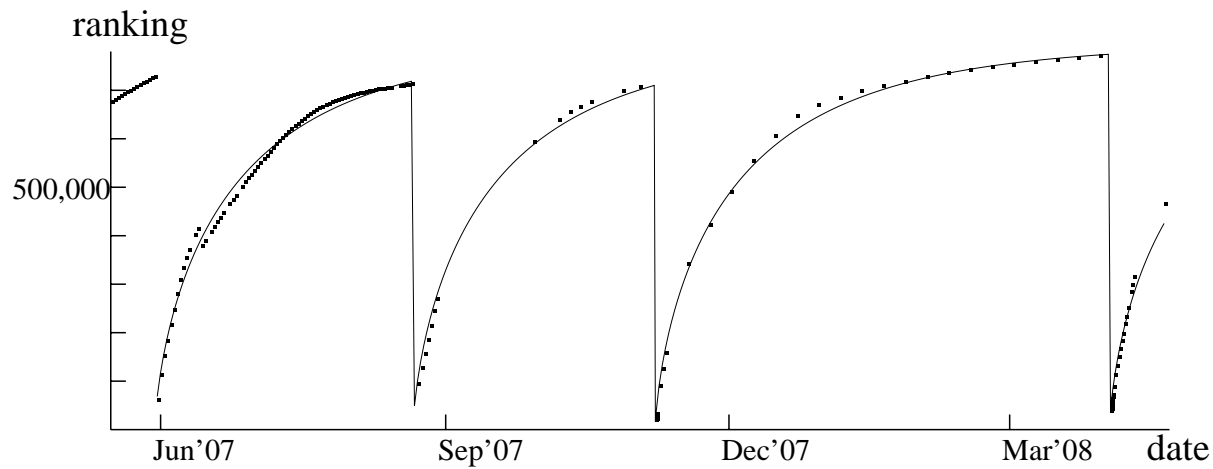


縦軸 $w_i^{(N)}$, 横軸 i ,

左図： $b < 1$ 格差大 右図： $b > 1$ 平等に近い

$b > 1$ のとき，低コスト ($c_2 \ll c_1$) ならば，非売れ筋の売上合計が無視できない（ロングテールビジネスモデル成立の可能性）

アマゾン is ロングテールに非ず



3パラメータで98点のデータを当てはめ

$$(N^*, a^*, b^*) = (8 \times 10^5, 6 \times 10^{-4}, 0.81)$$

- $b < 1$ 「不平等」上位 $o(N)$ 冊が売り上げの全てを占める

アマゾン is ロングテールに非ず

ロングテール型ではなく、ビッグヒット依存型のビジネスモデル

データに関する要約と予告

Amazon.co.jpのランキング, 2ch.netのスレッド一覧への当てはめ
2chは次ページ以降

- 「Amazonはロングテールに非ず」

現場や詳しい人は気づいていたのでは？

- ・部外秘に属するだろう情報を, 公開情報だけで実証する方法を与えた

- 活動の昼夜差の抽出

- N の時間変化(新刊本, 絶版)も $w_i(t)$ に織り込み可能

- 先頭注目効果も数学理論はできた(位置依存性)

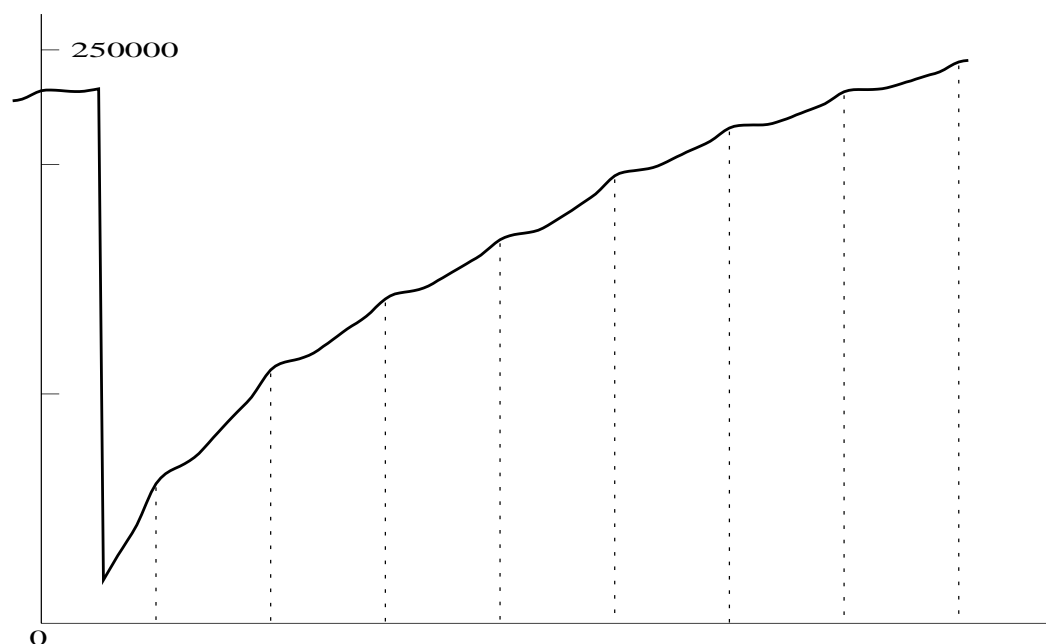
アマゾン書店:

更新が1時間に1度, 書店がアルゴリズムを決められる(上位には影響)

2ch.netスレッド一覧:

$N = O(10^3)$ 小さい 揺らぎ, 関心の集中「祭り」(ポワッソンが悪い)

社会活動の昼夜差



Amazon.co.jp
(毎時のランキングデータ自動収集)

昼夜差
(深夜から早朝まで**全体**
の活動停滞)

w_i が時刻依存性を持つ場合を (上手に) 適用して当てはめる

仮定 : 強度の共通時刻依存と日周期

毎日定時のデータは一様な場合に帰着

(Pareto 分布やロングテールの結論は相対強度の分布の意味で成立)

ジャンプ率の時刻依存性

$$X_i^{(N)}(t) \doteq N Y_C(O, t) = N - N \int_W e^{-\int_0^t w(s) ds} \lambda(dw)$$

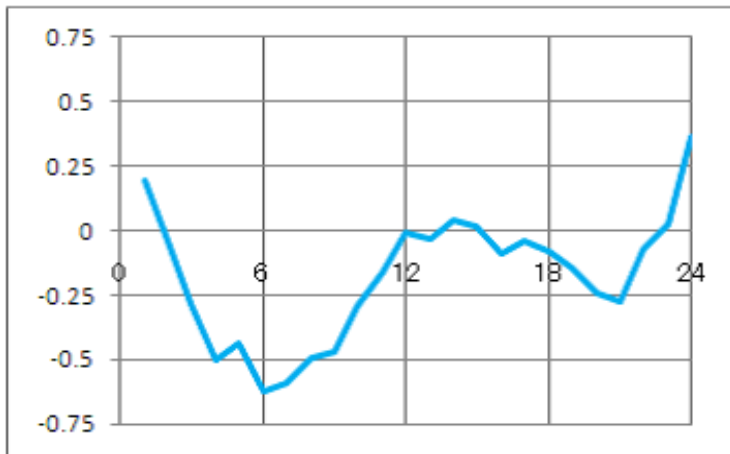
強度の共通時刻依存と日周期の仮定：

$W = \{w(t) = w A'(t) \mid w > 0\}$, for fixed periodic θ and $A = t + \theta(t)$

$$X_i^{(N)}(t) \doteq N - N \int_0^\infty e^{-w A(t)} \lambda(dw)$$

毎日定時のデータは一様な場合に帰着 $X_i^{(N)}(t) \sim N y_C, \text{一様}(A(t))$
(Pareto 分布やロングテールの結論は相対強度の分布の意味で成立)

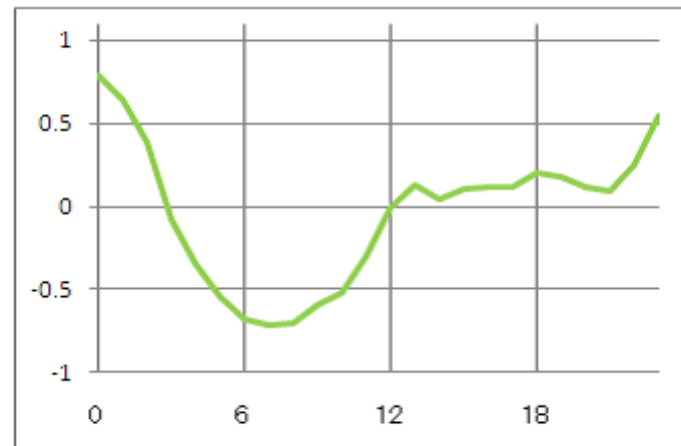
昼夜差 (Amazon.co.jp)



方法 1 (ジャンプ直後のデータ $aA(t) \ll 1$)

$$x_C(t) \simeq N\Gamma(1-b)a^b A(t)^b + O(A(t))$$

24時間間隔で2重に比をとって N, b, a を消去



方法 2 (テール付近のデータ)

順位の平均時間変化を線形近似

- 22–00時に購入活動活発, 03–09時に不活発 (注. 1.5時間前)
- 2ch.netと整合 (後述)
- 帰宅後就寝前型 (仕事時間型や昼夜逆転型ではない)

課題 (Amazon.co.jp)

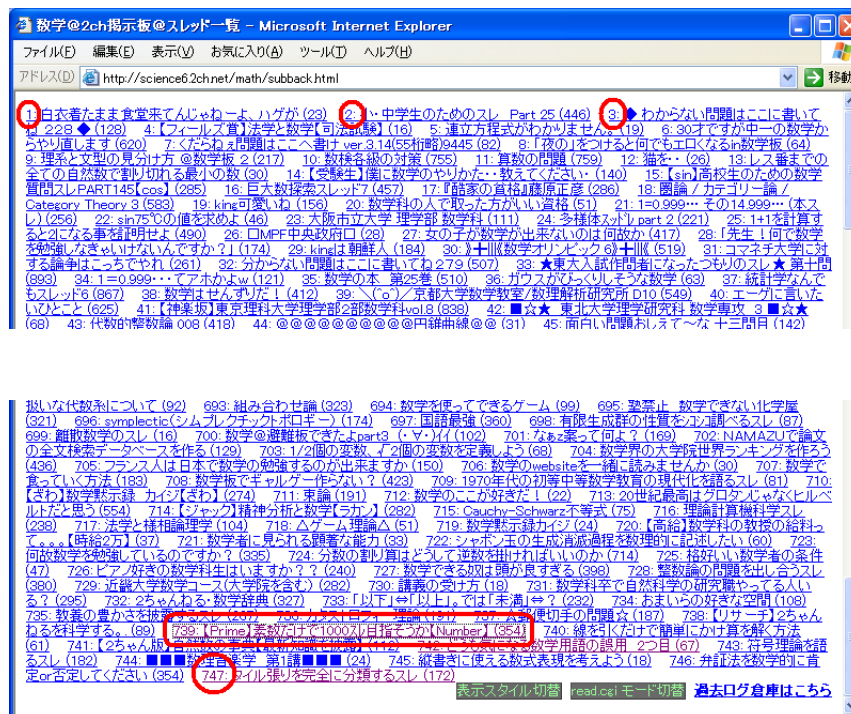
- ・ 相対強度の分布（「アマゾン」はロングテールに非ず）の結論はOK
- ・ 一歩進んで、強度の時刻依存性（**昼夜差**）も定量的に測りたい。**難**
 - アマゾン書店は1時間に一度の更新
 - 多数の本の順位を同時観測すれば方法があるが、順位一覧のページは無いのでサーバに負担（**後述**）

2ch.netのスレッド一覧

2ちゃんねる：
web 掲示板の巨大な集まり

スレッド（ページ）一覧：
書き込んだスレッドが1位
先頭に跳ぶ規則

（注：sage 進行は除く）

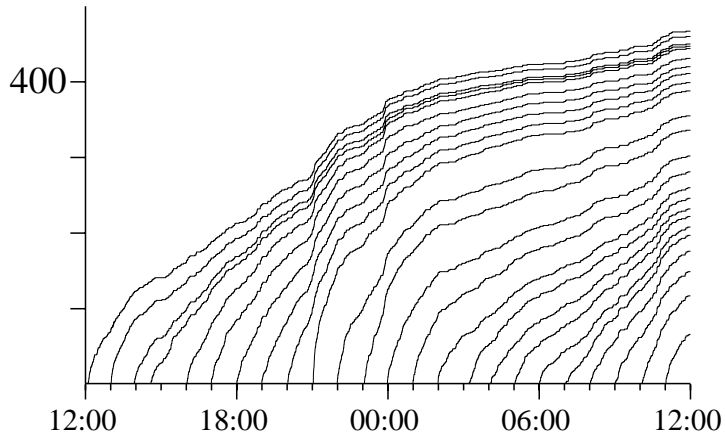


N は 10^3 程度で小さい（ので、数%の揺らぎがある）が、
先頭に跳ぶ規則が明らかなので、踏み込んだ検証が可能
スレッド一覧の1ページに全スレッドの順位情報 全数調査可能

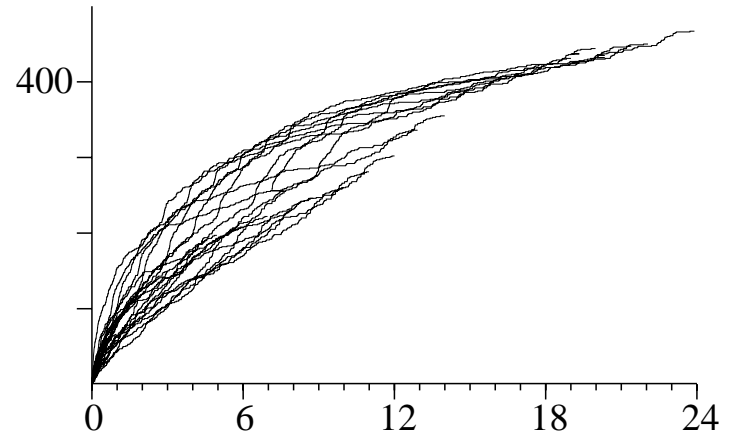
スレッド一覧の順位変化

竹島佑介君 (2008年度東北大M2, 現富国生命)

自動化によって全粒子 ($N \sim 700$) の24時間順位変化データ



24個のスレッド, 最後に書き込まれて以降の順位変化



最後に書き込まれた時刻を0に取り直して重ねた図

強度の時刻依存性ため, 1つの関数 $y_C(t)$ に重ならない

累積ジャンプ数

時刻依存性 $A(t)$ についてのデータ処理

累積総ジャンプ数 $S^{(N)}(t) = \sum_{i=1}^N \int_0^t w_i^{(N)}(s) ds$ で時間を計る

共通時刻依存性 $\int_0^t w_i^{(N)}(s) ds = w_i^{(N)} A(t)$

$S^{(N)}(t) \sim A(t)Z(N); Z(N) = \sum_{i=1}^N w_i^{(N)}$ (大数の法則)

$A(t)$ $S^{(N)}(t)$ 時刻依存性

λ $X_i^{(N)}(t)$ (相対)ジャンプ率分布

強度の共通の時刻依存性と累積書込数

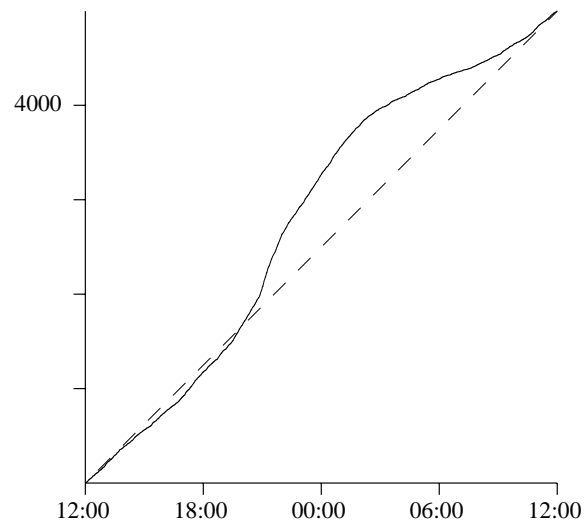
(再掲) 共通の時刻依存性の仮定:

$$\int_0^t w_i^{(N)}(s) ds = w_i^{(N)} A(t)$$

累積総書込数

$$S^{(N)}(t) = \sum_{i=1}^N \int_0^t w_i^{(N)}(s) ds \sim A(t)Z(N);$$

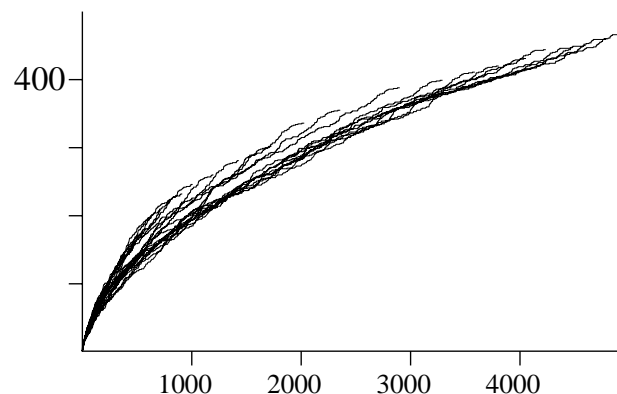
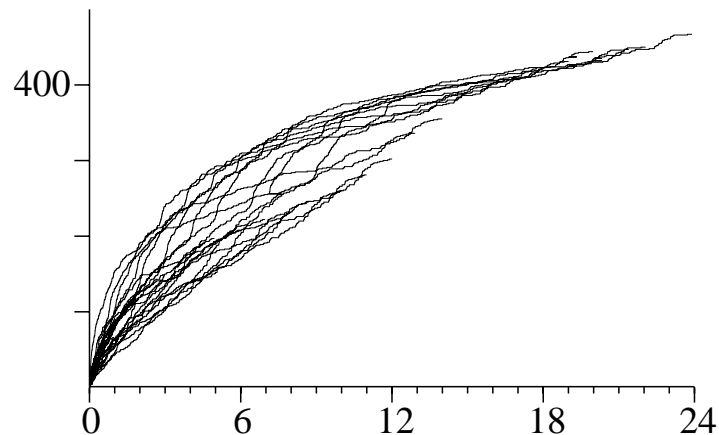
$$Z(N) = \sum_{i=1}^N w_i^{(N)}$$



集計 小林孝長君 (2009年度東北大M2, 現仙台二高)

共通の時刻依存性の検証と Pareto 指数の普遍性

(再掲) $X_i^{(N)}(t) \sim N(1 - \int_0^\infty e^{-wS^{(N)}(t)} \lambda(dw)); \lambda([w, \infty)) = (\frac{a}{w})^b, w \geq a$



(再掲) 横軸実時刻

横軸累積総書込数 (時間変更)

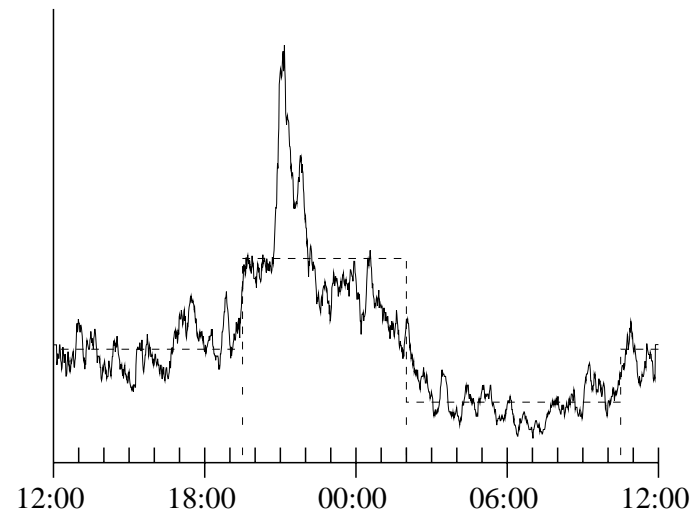
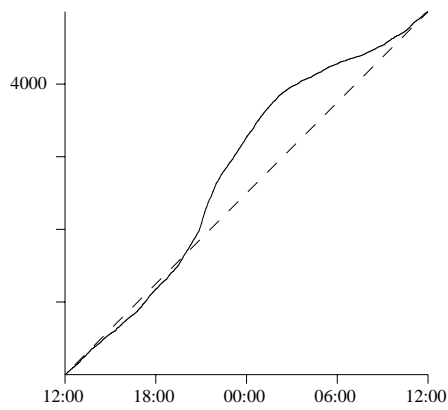
- 共通の時刻依存性は実用的に良さそうである

$X_i^{(N)}(A^{-1}(t)) (A(t) \sim S^{(N)}(t)) \quad b = 0.872 < 1$

- Amazon.co.jp も 2ch.net も共通して $b < 1$

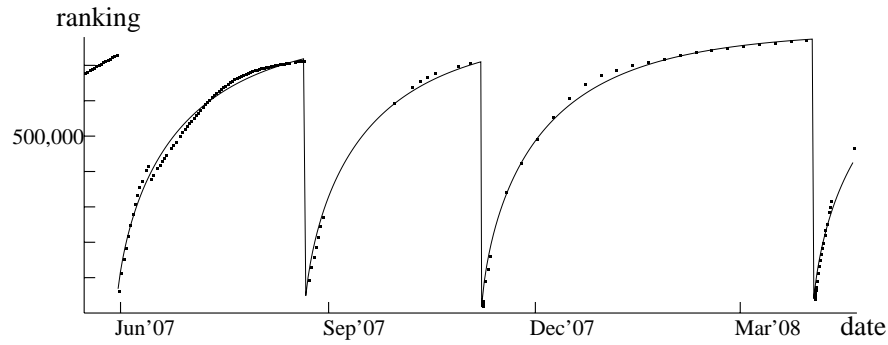
本の購入はベストセラーに, 書き込みは人気スレッドに**集中**

昼夜差

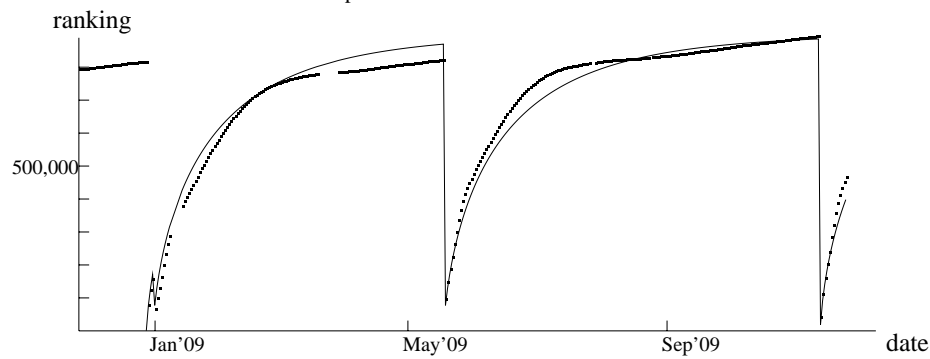


- 20-01時に書込活動活発，03-09時に不活発（2ちゃんねる）
（アマゾン書店と整合）

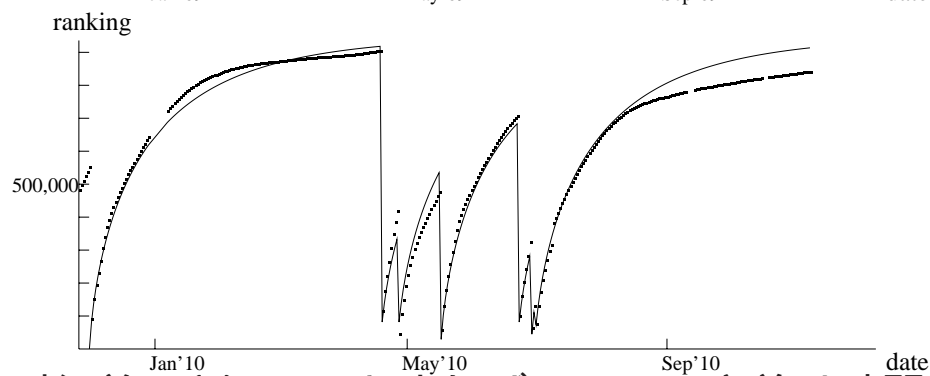
冊数の変化



$N = 80$ 万 (2007 年)



$N = 90$ 万 (2009 年)



$N = 95$ 万 (2010 年)

刊行前 $w(t) = 0$ とすればよい (昼夜差は時間変更). ライフサイクルの検出

全体の要約

- 先頭に跳ぶ規則の流体力学的極限の存在定理
 - 動きは巨視的には決定論的
 - 極限分布は蒸発項で動く inviscid Burgers 型方程式系
 - 時刻・位置依存性を入れられる（昼夜差，上位注目効果）
 - 特性曲線（1位に跳ぶ以外の順位変化）
- 流行度の順位付け（ランキング）のモデルとしての確率順位付けモデル
 - 流行度は確率的
 - 普遍性
 - ロングテールは順位付けが確率的であることが顕著な領域
- Amazon.co.jp のランキング，2ch.net のスレッド一覧への当てはめ
 - 「Amazon はロングテールに非ず」
 - 部外秘に属するだろう情報を公開情報だけで実証する方法
 - 活動の昼夜差の抽出

文献

服部哲弥「Amazon ランキングの謎を解く」化学同人, 2011.

服部哲弥, 日本OR学会誌 **57-6** (2012.6) 3-8.



K. Hattori, T. Hattori, Stochastic Processes Appl. **119** (2009)

K. Hattori, T. Hattori, Funkcialaj Ekvacioj **52** (2009)

K. Hattori, T. Hattori, RIMS Kokyuroku Bessatsu **B21** (2010)

Y. Hariya, K. Hattori, T. Hattori, Y. Nagahata, Y. Takeshima, T. Kobayashi, Tohoku Mathematical Journal **63-1** (2011)

T. Hattori, S. Kusuoka, Latin American Journal of Probability and Mathematical Statistics (2013)

Google 検索キーワード 服部哲弥