確率的順位付け

針谷祐　（東北大・理）
服部久美子　（首都大・数理）
服部哲弥　（慶應大・経済）
永幡幸生　（大阪大・基礎工）
竹島佑介　（富国生命）
小林孝長　（仙台二高）

2010.11.20 語ろう「数理解析」（芝浦工業大学大学院 工学研究科）

START

1．確率的順位付け模型

数学会応用数学分科会一般講演（近畿大学 08'03）

Stochastic ranking process (**確率的順位付け模型**)

N **自然数（粒子数），時刻** $t \geqq 0$，

$X^{(N)} = (X_1^{(N)}, \cdots, X_N^{(N)})$：$\mathbb{N}^N$ **値確率過程**

初期値 $X_i^{(N)}(0) = x_i^{(N)}$; $\{x_i^{(N)} \mid i = 1, 2, \cdots, N\} = \{1, 2, \cdots, N\}$

$$X_i^{(N)}(t) = x_i^{(N)} + \sum_{k=1}^{N} \int_0^t \mathbf{1}_{X_k^{(N)}(s-0) > X_i^{(N)}(s-0)}\, \nu_k^{(N)}(ds) + \int_0^t (1 - X_i^{(N)}(s-0))\, \nu_i^{(N)}(ds),\ i = 1, \cdots, N,\ t \geqq 0$$

$\mathbf{1}_A$ **は事象** A **の定義関数，**

$\nu_i^{(N)}$: $(\Omega, \mathcal{B}(\mathbb{R}_+)) \to \mathbb{Z}_+$ **（増分１に集中した）ポワッソンランダム測度；**

i **について独立，強度** $\rho_i^{(N)} = \mathrm{E}[\,\nu_i^{(N)}\,]$ **は連続測度** $(\rho_i^{(N)}(\{t\}) = 0)$

例（一様）： $\rho_i^{(N)}((0, t]) = w_i^{(N)} t$

ポワッソンランダム測度（ポワッソン確率過程）

(数学的定義：　納得力ゼロの人に自発的再構成を期待する説明方法)

（増分１に集中した）強度 $\rho = \mathsf{E}[\,\nu\,]$ **のポワッソンランダム測度** $\nu : (\Omega, \mathcal{B}(\mathbb{R}_+)) \to \mathbb{Z}_+$

- **確率１のサンプル** $\omega \in \Omega$ **毎に** $\nu(\omega, \cdot)$ **は** $\mathbb{Z}_+$ **に値を取る** $\mathbb{R}_+$ **上の測度（**$\mathbb{R}_+$ **にばらまかれた個数）**
- $\nu(A) = \nu(\cdot, A)$ **の分布は平均** $\rho(A)$ **のポワッソン分布**
- $A \cap B = \emptyset$ **ならば** $\nu(A)$ **と** $\nu(B)$ **は独立**

$\rho((0,t]) = t$，$\geqq 0$，**のとき，** $Z_t = \nu((0,t])$ **は一様ポワッソン過程**

先頭に跳ぶ規則

時刻tの粒子iの順位（再掲）$X_i^{(N)}(t) = x_i^{(N)}$

$$+\sum_{k=1}^{N}\int_0^t \mathbf{1}_{X_k^{(N)}(s-0)>X_i^{(N)}(s-0)}\,\nu_k^{(N)}(ds) + \int_0^t (1 - X_i^{(N)}(s-0))\,\nu_i^{(N)}(ds)$$

添字i＝粒子名（ゼッケン），$X_i^{(N)}(t)$＝最後に先頭に跳んだ順に並べるときの時刻tでの列の中でのiの位置（順位）

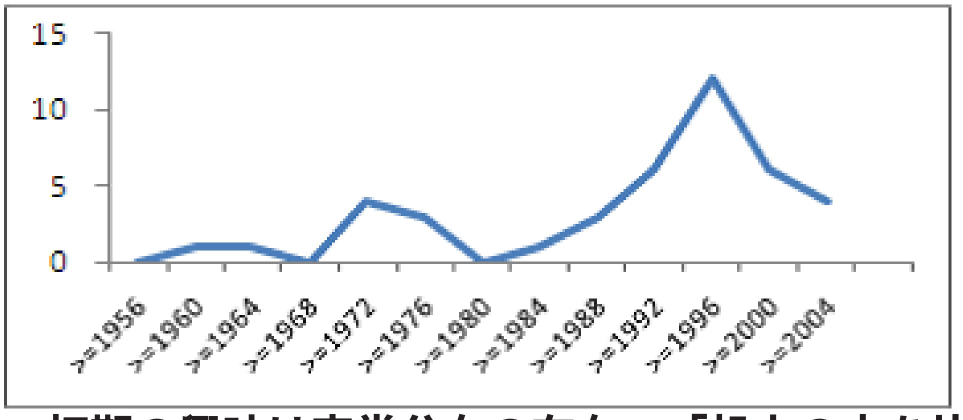
$\nu_i^{(N)}((a,b])$：$a < t \leqq b$に粒子iが先頭に跳んだ$(X_i^{(N)}(t) = 1)$回数

先頭に跳ぶ時刻$\nu_i^{(N)}(\omega) = \displaystyle\sum_{j=1}^{\infty} \delta_{\tau_{i,j}^{(N)}(\omega)}$；$0 < \tau_{i,1}^{(N)} < \tau_{i,2}^{(N)} < \cdots$

- 確率１で$\{\tau_{i,j}^{(N)} \mid i = 1,\cdots,N,\ j \in \mathbb{N}\}$は互いに異なる
- $\displaystyle\int F(s)\nu_i^{(N)}(ds) = \sum_j F(\tau_{i,j}^{(N)})$
- $X_i^{(N)}(\tau_{i,j}^{(N)}) = 1$
- $X_{i'}^{(N)}(\tau_{i,j}^{(N)}) = X_{i'}^{(N)}(\tau_{i,j}^{(N)}-) + 1\ (\forall i';\ X_{i'}^{(N)}(\tau_{i,j}^{(N)}-) < X_i^{(N)}(\tau_{i,j}^{(N)}-))$

確率ランキング模型への注

- 流行に従う順位付けの究極の単純化（超整理法）
- Move-to-front 規則の文献は M.L.Tsetlin (1963) 以来，再発見を繰り返す



- 初期の興味は定常分布の存在　「机上の本を片付けられるとがっかり」
- メモリーへのデータ配置 (Least-Recently-Used caching)
- 最近の流行 (?)：top-to-random shuffling　（時間反転系）

（ジャンプ率 w は一種，モンテカルロ法の定常分布への収束との関連）

- 誰も言及しない待ち行列との対応：右端がアトラクションの入場口．極端な入場制限のため，諦めて出て行くことだけで列が動く．出た分だけ後ろに並んで一定の長さ．ジャンプ率はあきらめの早さ．

- 確率ランキング模型：ジャンプ率分布の大数の法則（cf. 粒子 i の区別）
  流体力学的極限：微視粒子 (確率過程) 多数→巨視的分布 (PDE)

＊．目次

２．大数の法則（軌道，位置強度結合経験分布）

ジャンプ済み粒子と未ジャンプ粒子の境界：$Y_C^{(N)}(t) = \dfrac{1}{N}\sum_{i=1}^{N} \mathbf{1}_{\tau_{i,1}^{(N)} \leqq t}$

命題． $\lambda_t^{(N)} = \dfrac{1}{N}\sum_{i=1}^{N} \delta_{\rho_i^{(N)}((0,t])} \to \exists\lambda_t \ (N\to\infty)$ **ならば，**

$Y_C^{(N)}(t) \to y_C(t) = 1 - \int_0^\infty e^{-s}\lambda_t(ds) \ (N\to\infty$，**概収束**$)$

さらに λ_t **が** t **にで連続なら** path **として**（**時刻の各点収束位相で**）**概収束**◇

証明：　独立確率変数の和の大数の強法則：$Y_C^{(N)}(t) - \mathsf{E}[\,Y_C^{(N)}(t)\,] \to 0$
＋ポワッソン分布：$\mathsf{P}[\,\tau_1 > t\,] = \mathsf{P}[\,\nu((0,t]) = 0\,] = e^{-\rho((0,t])}$

流れの中の粒子の軌道

$x_i^{(N)} = 1$，$0 \leqq t < \tau_i^{(N)}$ **のとき，** $X_i^{(N)}(t) - 1$

$$= \sum_{k=1}^{N} \int_0^t \mathbf{1}_{X_k^{(N)}(s-0) > X_i^{(N)}(s-0)} \, \nu_k^{(N)}(ds) = \sum_{i=1}^{N} \mathbf{1}_{\tau_{i,1}^{(N)} \leqq t} = N Y_C^{(N)}(t)$$

注目粒子が先頭に跳んだ時刻を $t = 0$ **とすると**

次にジャンプするまでの粒子の軌道

命題（再掲） $\displaystyle \frac{1}{N} \sum_{i=1}^{N} \delta_{\rho_i^{(N)}}((0,t]) \to \lambda_t \ \Rightarrow \ Y_C^{(N)}(t) \to y_C(t) = 1 - \int_0^\infty e^{-s} \lambda_t(ds)$

- **強度が小さい（ジャンプ率** w **が小の）粒子のジャンプ数は揺らぐが，そういう粒子が多数あるので，軌道は決定論的**
- $y_C(t)$ **は観測され，強度（ジャンプ率）分布** λ **を推定できる（後述）**

位置ー強度結合経験分布の収束

強度 ρ（ジャンプ率）と規格化順位 $Y_i^{(N)} = \frac{1}{N}(X_i^{(N)} - 1)$ の結合経験分布：

$$\mu_t^{(N)} = \frac{1}{N}\sum_{i=1}^{N} \delta_{(\rho_i^{(N)}, Y_i^{(N)}(t))}$$

混合液の濃度分布をミクロに計測（分布値確率変数：度数分布，揺らぎあり）

定理． $\mu_0^{(N)} \to \exists\mu_0$ $(N \to \infty)$，かつ，各 $0 \leqq s < t$ に対して $N \to \infty$ で

$$\frac{1}{N}\sum_{i=1}^{N} \delta_{\rho_i^{(N)}((s,t])} \to \int_{\mathcal{M}(\mathbb{R}_+)} \delta_{\rho((s,t])} \Lambda(d\rho);\ \Lambda(d\rho) = \mu_0(d\rho \times [0,1))$$

ならば，各 $t > 0$ に対して $\mu_t^{(N)} \to \mu_t$ $(N \to \infty$，概収束）．
ここで μ_t は次頁にあらわに与える非ランダムな分布． ◇

μ_t **：混合液の濃度分布をマクロに計測（確定した分布：揺らぎなし）**

極限分布

$\mu_t^{(N)} \to \mu_t$．ここでμ_tは： $U(d\rho, y, t) := \mu_t(d\rho \times [y, 1))$

$$= \begin{cases} e^{-\rho((t-t_0(y,t),t])} \Lambda(d\rho), & 0 \leqq y \leqq y_C(t), \\ e^{-\rho((0,t])} U(d\rho, \widehat{y}(y,t), 0), & y_C(t) \leqq y < 1. \end{cases}$$

t_0は$y_A(t_0, t) = 1 - \displaystyle\int_{\mathcal{M}(\mathbb{R}_+)} e^{-\rho((t-t_0,t])} \Lambda(d\rho)$，

$\widehat{y}$は$y_B(y, t) = 1 - \displaystyle\int_{\mathcal{M}(\mathbb{R}_+)} e^{-\rho((0,t])} \mu_0(d\rho \times [y, 1))$の，逆関数

$y_C(t) = y_A(t, t) = y_B(0, t)$

極限を記述する偏微分方程式

簡単のため，有限種で密度（ジャンプ率関数）を持つ強度の場合：

$$\Lambda = \sum_\beta r_\beta \delta_{\rho_\beta},\ \rho_\beta(A) = \int_A w_\beta(u)\,du,\ w_\beta, r_\beta > 0$$

定理． $\sum_\beta r_\beta = 1$, $\sum_\beta r_\beta w_\beta(t) < \infty$, $u_\alpha :\ [0,1) \to \mathbb{R}_+$: **非負滑らか狭義減少，** $\sum_\beta u_\beta(y) = 1 - y$ **ならば，** $U_\alpha(y,t) = U(\{\rho_\alpha\}, y, t)$ **は次の偏微分方程式系の初期値問題の時間大局的一意古典解：**

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

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

- **解が逆関数** $(t_0,\ \widehat{y})$ **で書かれる理由 － 特性曲線の方法** (y_A, y_B, y_C)
- **1次元非圧縮性混合流体の蒸発による運動における時刻** t **に** y **の右にある流体** α **の量**

定理の意味　ー　簡単なこと

- ミクロはランダム多自由度，マクロは決定論的少自由度
  (cf. Amazon.co.jp の本 $N = O(10^6)$)
- 順位の先頭付近は jump 率の高い粒子が多く tail は低い粒子が多い
- 上位側 $(y \leqq y_C(t))$ では初期値 $(\mu_0 = U(\cdot,\cdot,0))$ によらない

ー randomness は先頭へのジャンプのみで，それが位置によらない：平均場的

- cf.　TASEP（1 次元格子上の多粒子系で空き格子に左から粒子が移れる）の流体力学的極限は inviscid Burgers　ー　解が爆発

さっきの定理：位置ジャンプ率結合分布の PDE は時間大局的古典解
（確率ランキング模型はやさしい）

定理の意味　−　ロングテールビジネスモデル

- **簡単のため，時間的に一様なジャンプ率で定常** $(t=\infty)$ **な場合：**

$\rho((0,t]) = wt$，w **の分布** $\lambda(dw) = \mu_\infty(dw \times [0,1))$，

$$\mu_\infty(dw \times [y,1)) = e^{-wt_0(y)}\lambda(dw); \quad 1-y = \int_{\mathbb{R}_+} e^{-wt_0(y)}\lambda(dw)$$

- **ロングテール（列の後ろ＝その他大勢＝大多数の，めったに先頭になれない粒子）**

時刻 t **以降最初に先頭に飛んだ粒子** $I(t)$ **の直前の位置：** $C_N(t) = X^{(N)}_{I^{(N)}(t)}(t)$

$[y,1)$ **（**tail **側）にいたものの割合** $\mathrm{P}[\,\frac{1}{N}C_N(t) \geqq y\,] \;\to\; \dfrac{\int_{\mathbb{R}_+} w\,\mu_\infty(dw\times[y,1))}{\int_{\mathbb{R}_+} w\,\lambda(dw)}$

- Amazon.co.jp **ランキングではテール側からの売り上げへの寄与**

LRU-caching **では** cache fault **の確率**

例．Pareto **分布：** $a,b>0$, $\lambda([0,w]) = 1-(a/w)^b$ $(w \geqq a)$ $(w_i = a(N/i)^{1/b})$

$y>0$ **のとき分子は有限正**

$b>1$**：** $\displaystyle\int_{\mathbb{R}_+} w\,\lambda(dw) = \frac{ab}{b-1}$, $0<b<1$**：** $=\infty$ **（**$y=+0$ **の寄与が発散）**

ロングテールビジネスが意味を持ちうるのは $b>1$

証明についてのコメント

強度が有限種類（Λ が有限集合）ならば各種類毎の収束→初期の結果に帰着
Λ が「巨大」（関数形が非可算無限種類）なとき：
主張の内容は，確率 1 の事象 $\tilde\Omega$ の sampleω について，任意の有界連続な f に対して

$$\lim_{N\to\infty}\int_{\mathcal{M}(\mathbb{R}_+)\times[0,1)} f(\rho,y)\mu_t^{(N)}(d\rho\times dy) = \int_{\mathcal{M}(\mathbb{R}_+)\times[0,1)} f(\rho,y)\mu_t(d\rho\times dy)$$

- $\mathcal{M}(\mathbb{R}_+)\times[0,1)$ **が可分完備距離空間なので，可算個の $f=f_n$ で言えば良いことが知られている →個別に Ω'_n を選んで証明し，$\tilde\Omega=\bigcap_n\Omega'_n$．　y 方向については $\chi[y,1)$ の重ね合わせで近似　→**
- $\displaystyle\lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N} g(\rho_i^{(N)})\,\mathbf{1}_{Y_i^{(N)}(\omega)(t)\geqq y} = \int_{\mathcal{M}(\mathbb{R}_+)} g(\rho)\mu_t(d\rho\times[y,1))$ **に帰着**
- **従属確率変数の大数の法則（右側の粒子が飛んだときだけ影響）　→**

同じ極限を持つ扱いやすい確率変数

補題． $y\leqq y_C(t)$ **のとき** $\displaystyle\frac{1}{N}\sum_{i=1}^{N}\left|\mathbf{1}_{Y_i^{(N)}(t)<y}-\mathbf{1}_{\nu_i^{(N)}((t-t_0(y,t),t])>0}\right|\to 0$, a.s.

最近の進展．（$\sharp\Lambda<\infty$ のとき） $\mu_t^{(N)}(\{w\}\times[y,1))$ **と極限の差を，**martingale**（＋小さい量）で表して，**Doob**の不等式と**Gronwall**の定理で証明できる．**Tagged particle**の収束も同様に証明（永幡幸生）**

３．Amazonデータへの当てはめとロングテール

- Amazon.co.jp ランキング：インターネット時代の新しい応用
- ロングテールビジネスモデルの検証
- 掲示板集合体2ch.netのスレッド一覧

ブログ集合体(ameblo.jp)，機関リポジトリの論文アクセス

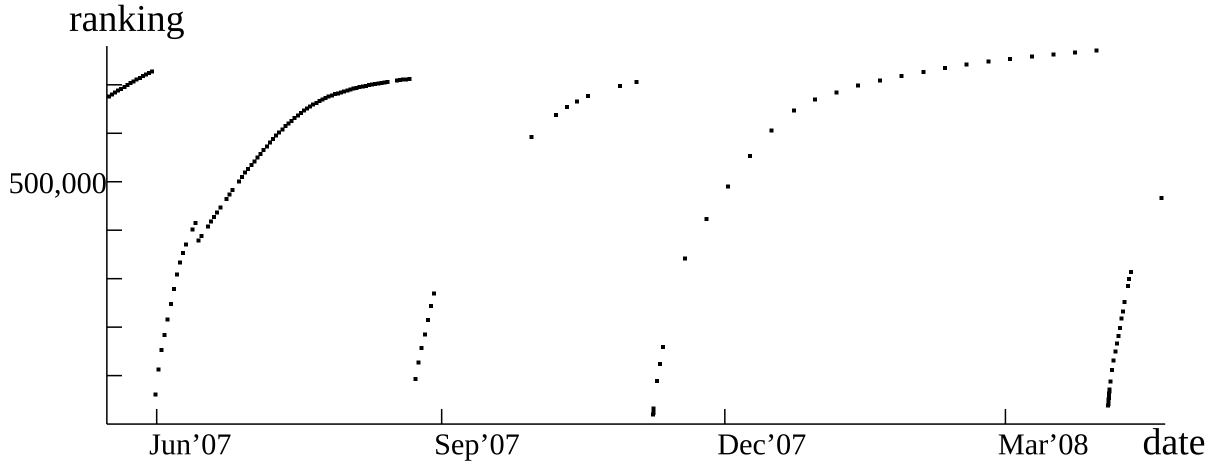
Amazon.co.jp ランキング

「どういう計算方法なのかわからないAmazonの謎順位」 (2006年末，佐々木将人（函館簡裁）@数理科学ML)

ランキングの時間変化

- （さほど売れない本を）しつこく観察すればわかる

実際は，モデルにたどりついて，公式を作ってから，検証のためにしつこく観察した



上位への大きなジャンプ　－　アマゾンのウェブで注文することに対応
１冊でも，中古でも

☆「（１冊買うだけで）１位に飛ぶか？」→大多数の本にとっては良い近似！

大多数の本にとって，最近の流行＝最後に売れた時刻　→
注文が独立ならば確率ランキング模型が良い近似

理論軌道への当てはめとPareto分布

$x_i^{(N)} = 1$，$0 \leqq t < \tau_i^{(N)}$ のとき $Y_i^{(N)}(t) = Y_C^{(N)}(t) + \dfrac{1}{N}$

本の順位を最後に売れた時刻を $t = 0$ と取り直して，次に売れるまでを $NY_C(t)$ に当てはめる

$$x_C(t) \simeq N - N \int_0^\infty e^{-s} \lambda_t(ds)$$

実際のデータに当てはめるためには，ジャンプ率の分布 λ_t が必要

社会学や経済学では（一般化された）Pareto 分布がよく用いられる

離散版（Zipf**の法則**）$w_i^{(N)} = a \left(\dfrac{N}{i}\right)^{1/b}$；$a$: **最低収入**，$b$: **平等性の指数**

$x_C(t) \simeq N\, y_C(t) \simeq N(1 - b(at)^b \Gamma(-b, at))$；$\Gamma(z,p) = \int_p^\infty e^{-x} x^{z-1}\, dx$

N, a, b **を与えれば決まる（データを使って統計的当てはめ）**

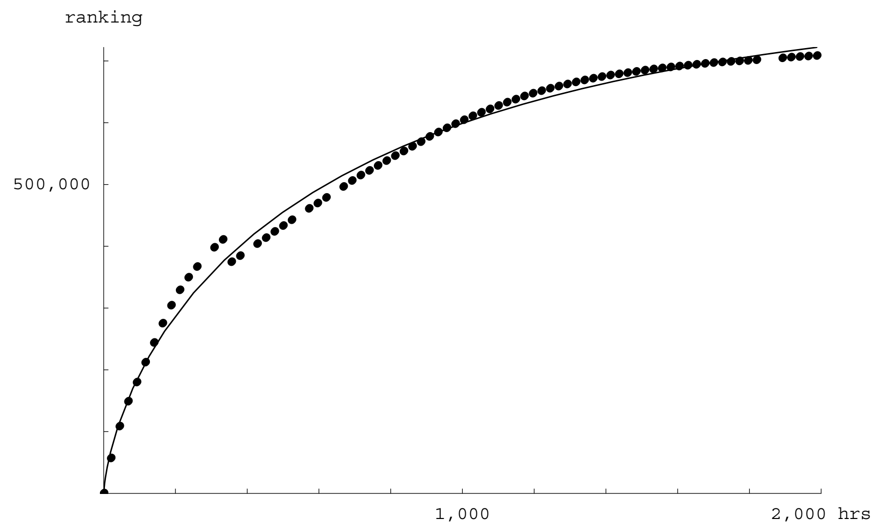
Amazon.co.jpはロングテールではない！

毎日21時のランキング

'07.6–9 **（逃したら諦める）**

$(N^*,\ a^*,\ b^*) =$

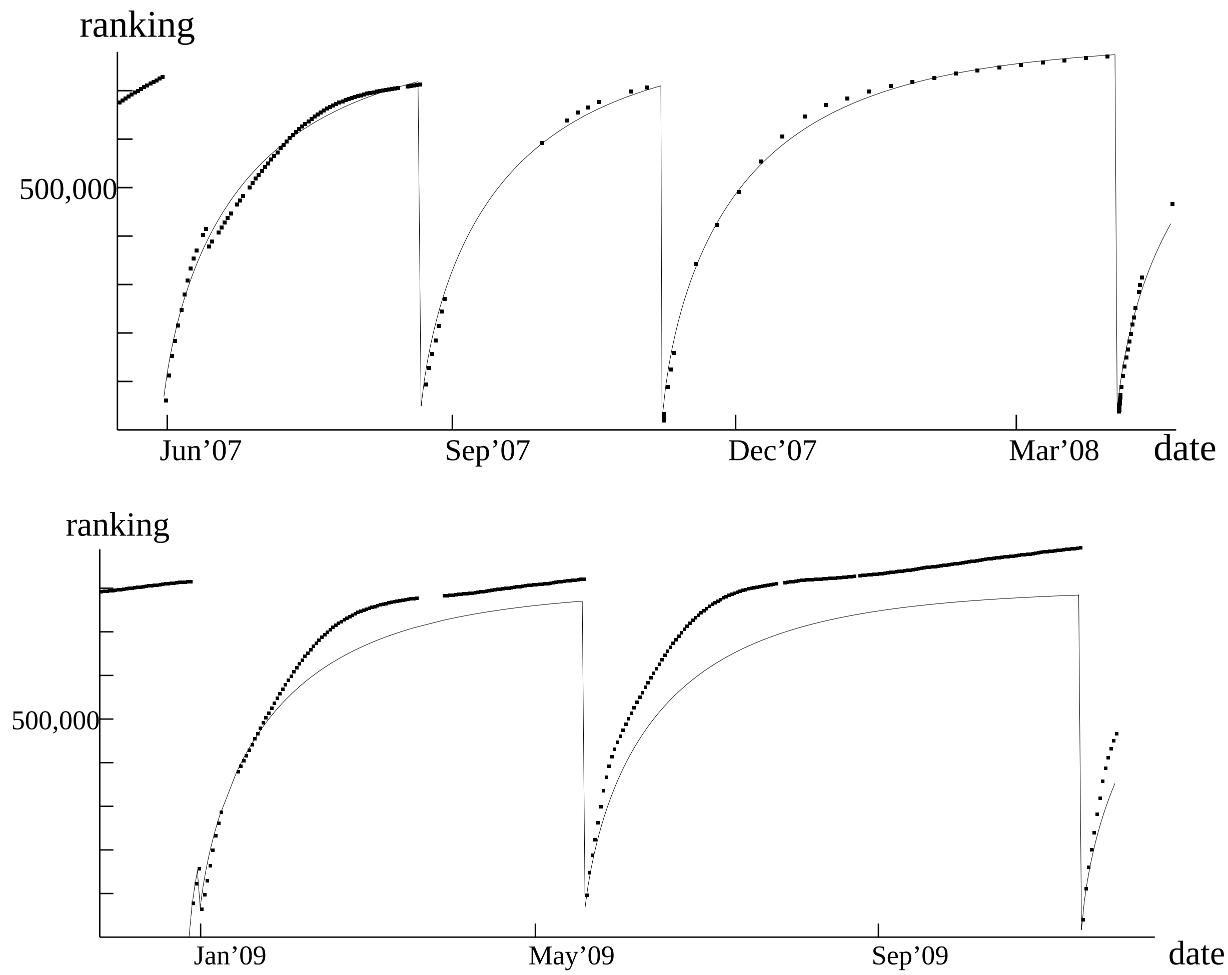
$(8\times10^5,\ 6\times10^{-4},\ 0.81)$



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

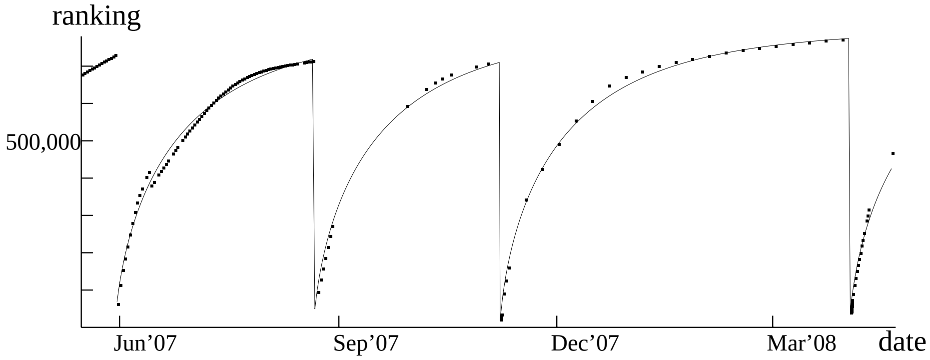
ロングテール型ではなく，ベストセラー依存型のビジネスモデル

4．2ch.netデータと指数の普遍性，強度の昼夜差

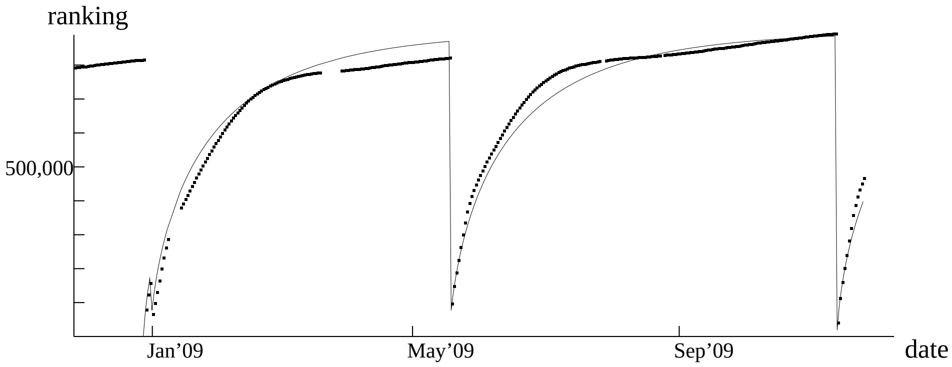


Amazon.co.jp ランキング（本の順位）　１時間毎に更新

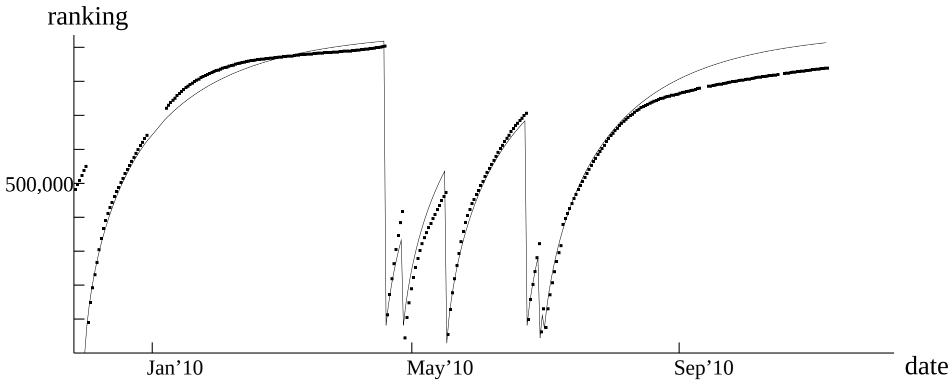
冊数の変化



$N = 80$ 万 (2007 年)

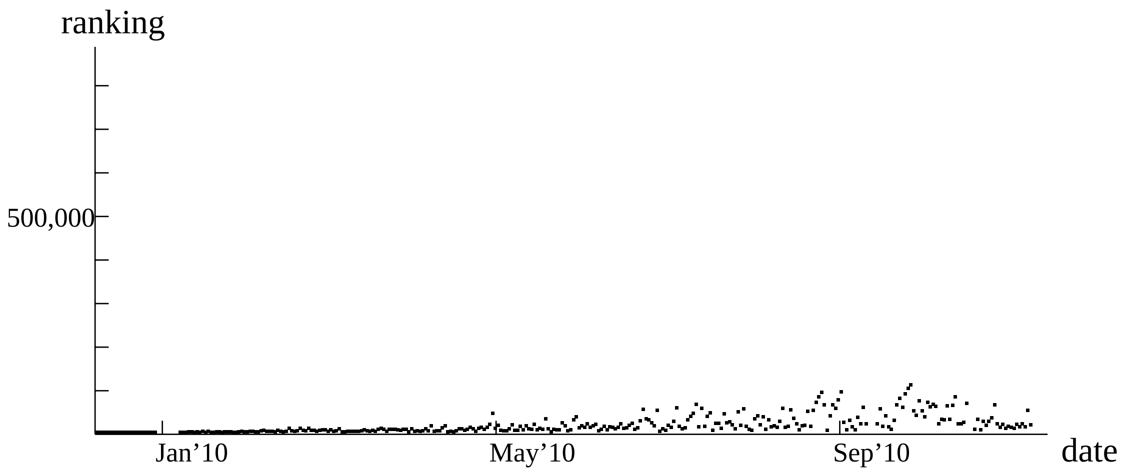


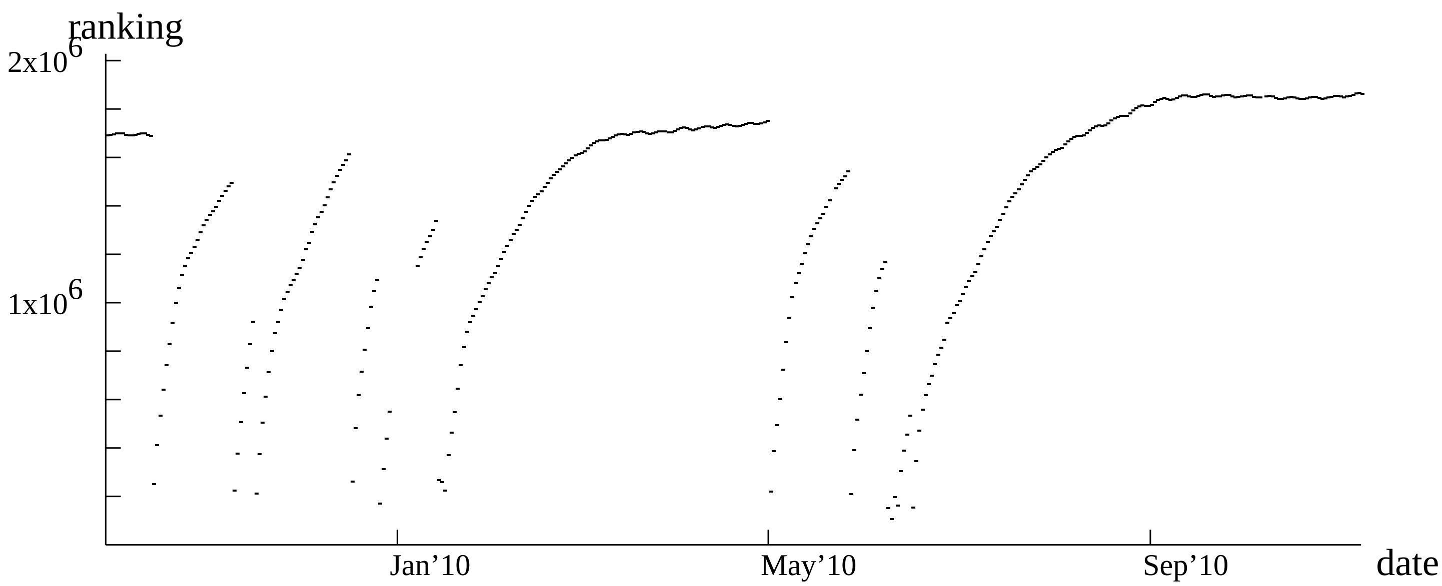
$N = 90$ 万 (2008 年)



$N = 95$ 万 (2009 年)

おまけ（有名文化人著作とamazon.com）





ジャンプ率の時刻依存性

社会活動の昼夜差

- ジャンプ率 $w_i^{(N)}$ が定数とした公式に当てはめて得た結論の再検証

→結論：正しいデータ利用方法だったので大丈夫

- データと拡張した理論から，購入活動の昼夜差を逆算できるか？

日内変動に影響されないデータ採取

λ_t: 時刻の関数丸ごと（無限パラメータ）はデータから決まらない

- **共通の時刻依存性**：$w_i^{(N)}(t) = w_i^{(N)} a(t),\ w_i^{(N)} > 0$

$$\lambda_t^{(N)} = \frac{1}{N} \sum_{i=1}^{N} \delta_{w_i^{(N)} A(t)} \to \exists \lambda_t;\ A(t) = \int_0^t a(u) du$$

- $a(t)$ が周期 24 の**周期関数**の場合：積分 $A(t) = t +$ **周期関数**
- 毎日**定時**に得たデータのみ用る（**逃したら諦める**）場合は，日内変動があっても平均ジャンプ率と等しい時間発展（変動効果は時刻の原点に吸収される）

今までのデータ解析は正当化された（共通でない時刻依存性は誤差扱い）

2ch.netのスレッド一覧

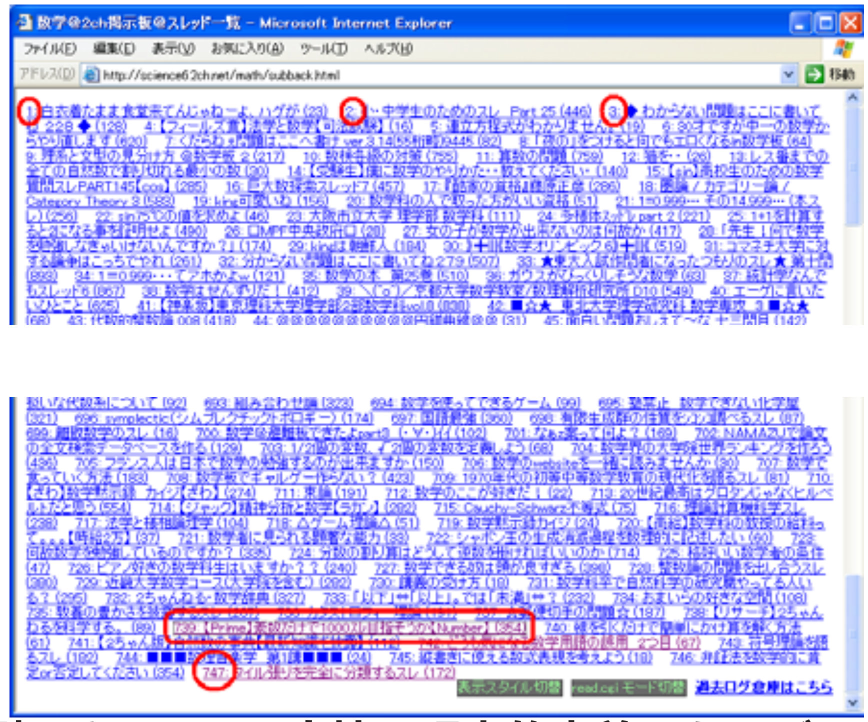
２チャンネル（2ch.net）**：**
web **掲示板の巨大な集まり**

スレッド（ページ）一覧：
「書き込んだスレが１位」：

一覧の順序は

move-to-front 規則

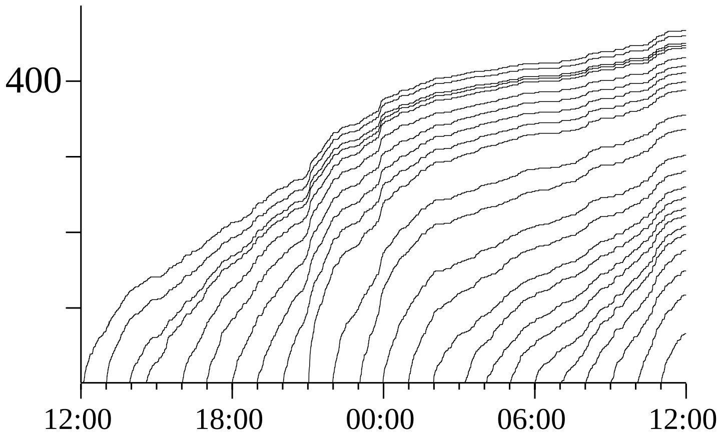
（注：sage 進行は除く）

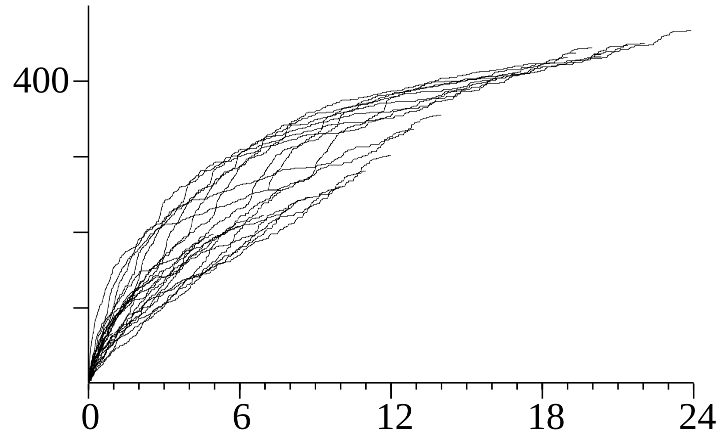


書込み情報公開　⇒　λ_t を直接観測できるので，直接の現実的意義はないが，方法論の検証やパラメータの普遍性の検討に使える

- **アマゾンは教えてくれないので，ランキングから探る実用上の意義**

スレ一覧の順位変化





スレたちの順位の時間変化
各スレ最後に1位になって（誰かが書き込んで）以降の時間変化
（全 $N = 697$ スレ，08/10/18 24時間約4000jump．24スレッドのみ表示）

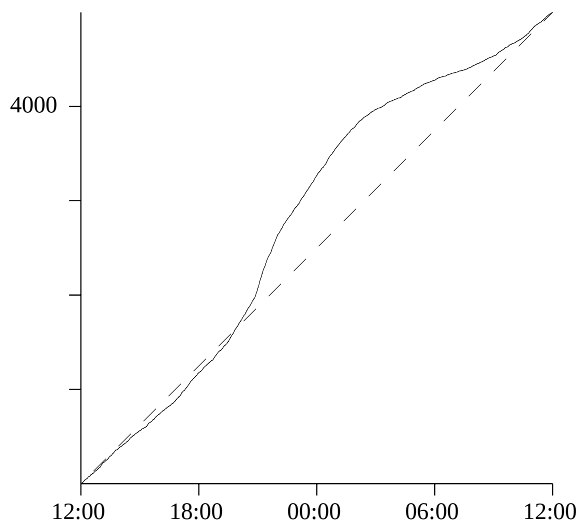
右図は1位になった時刻を0に取り直して重ねた図
昼夜差のため，1つの関数 $y_C(t)$ に重ならない
データ自動収集プログラム竹島佑介君（2008年度東北大M2，現富国生命）

共通の時刻依存性の抽出

- **共通の時刻依存性**：$A(t) = \int_0^t a(u)du;\ w_i^{(N)}(t) = w_i^{(N)} a(t) > 0$

累積総書き込み数 $S^{(N)}(t) \simeq \sum_{i=1}^{n} \rho_i^{(N)}((0,t]) = A(t)\,Z(N)$ **でスケール**

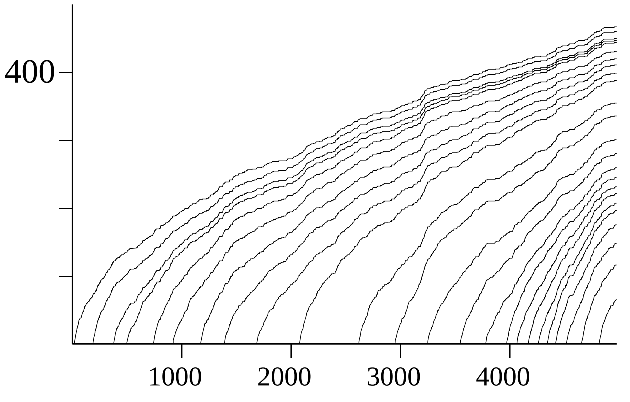
$$y_C(t) \simeq 1 - \int_{\mathbb{R}_+} e^{-w\,S^{(N)}(t)/Z(N)}\,\lambda(dw);\ Z(N) = \sum_{i=1}^{N} w_i^{(N)}$$

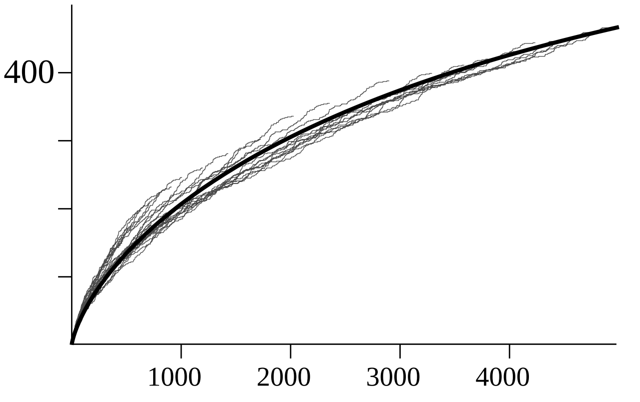


集計 小林孝長君

(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





実線：$y_C(A^{-1}t)$，Pareto分布，$b = 0.872$

- 共通の時刻依存性は実用的に良さそうである
- 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集中，書き込みは人気スレッドに集中

Amazon.co.jpはロングテールビジネスか？

先行研究：経営学的研究

Chevalier, Goolsbee $b = 1.2$ (Online bookstore の価格弾力性)
Brynjolfsson, Hu, Smith $b = 1.148$ (consumer welfareの評価)

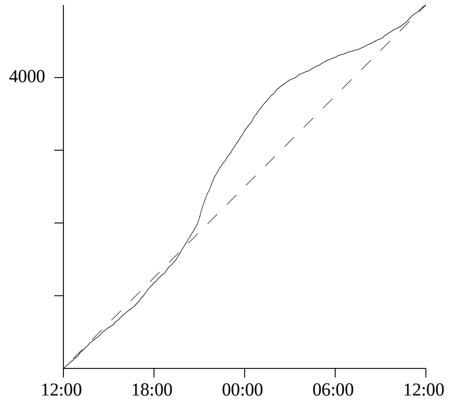
選んだ本の期間あたりの平均販売量を多数の本にわたってとる原始的方法

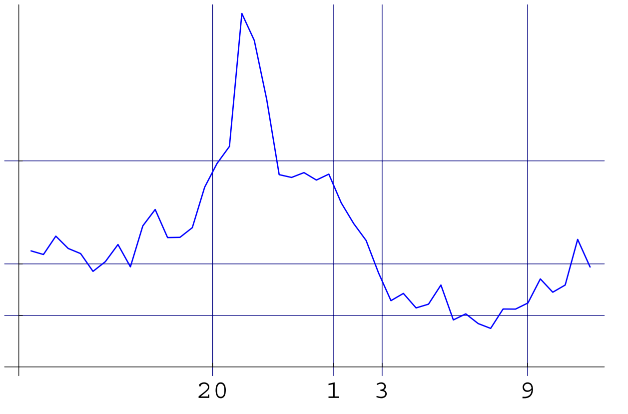
我々の結果と矛盾ー Online retailの経済効果は先行主張ほどはない？

インターネットを生かしたロングテール型リテールの草分けとして有名なアマゾン書店は，ロングテールビジネスモデルではない．ロングテールの「成功例」として C. Anderson, 'The Long tail' **などの記事という宣伝費無料の大規模宣伝を勝ち取り，実際は古典的大ヒットビジネスで利益を上げている，と思われる．**

傍証： Amazon **が売上げ詳細を隠すと言われている…**

昼夜差について（2ch.netデータ）





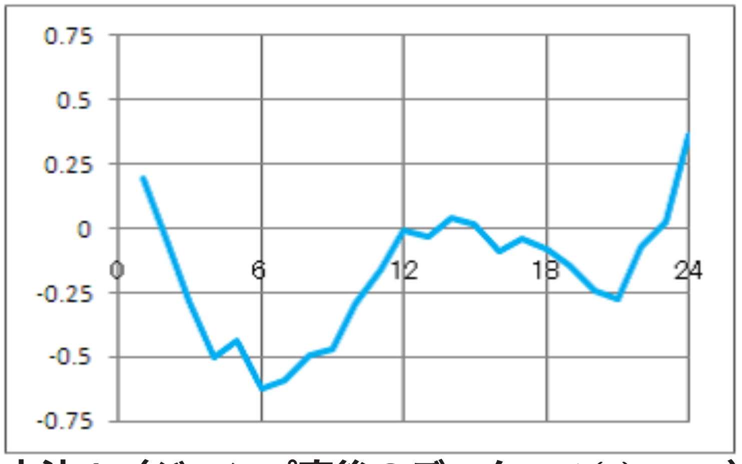
- 20–01時に書込活動活発，03–09時に不活発

Amazon.co.jpデータでも$O(10^2)$の本の総ジャンプ数を数えれば原理的に可能

1時間単位なので，1.5日/1冊(10万位)–9日/1冊(30万位)の本を集める必要

(cf. 10^3位=20分/冊, 10^2位=1分/冊)

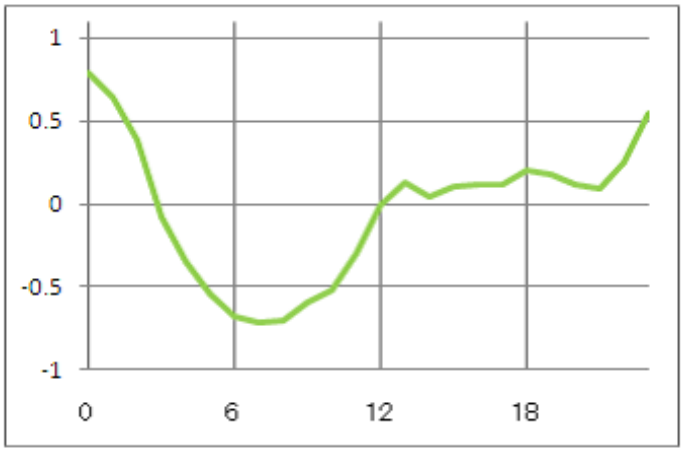
昼夜差について（Amazon.co.jp データ）



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 **と矛盾はない**
- **夜更かし型であって仕事時間型ではない**

5．まとめと先の話

- Stochastic ranking process (move-to-front 規則の流体力学的極限)
  - 巨視的な量（分布）に注目した無限粒子極限
  - 偏微分方程式の解に基づく収束定理
  - 巨視的視点からの自然な拡張（時刻依存性を持つジャンプ率）
- Amazon.co.jp データへの応用と 2ch.net データによる検証
  - 逆ラプラス変換によるジャンプ率（売れ行き）分布の統計的推定
  - Amazon.co.jp の顧客も 2ch.net の書き手も深夜活動し朝方寝ている
  - Amazon.co.jp も 2ch.net も上位人気対象に活動が集中
  - 数学が現在の応用の役に立たないというのは誤りである

Stochastic ranking processのこれから

- 流体力学極限の技術(martingale)による証明の整備と結果の精密化（永幡幸生）
- 空間位置に依存するジャンプ率
  - 偏微分方程式の自然な一般化（係数の位置依存性）
  - ランキングの売れ行きへのfeed back (local time)
- 個別の本のライフサイクル（栄枯盛衰）と定常な場合の差（検出問題）
- 微分方程式(macro)を極限に持つ粒子系(micro)の存在問題

  （物質に関する巨視的現象を分子運動として説明できるか？）

ランキングへの応用の展望

- 2ch.net 自動収集データによる位置ジャンプ率分布の極限定理の検証
- Amazon.co.jp データからの日周期の定量的導出
- 総粒子数の緩やかな増加（新規出版年８万冊）のモデルへの組み込み
- 機関リポジトリの論文アクセス（アクセスの集中）
- データを逆ラプラス変換して分布を統計的に推定する一般原理

Stochastic ranking processを離れた方向

（背後の社会現象の理解が進めば，データによるranking process**の検証も進む）**

- Pareto**分布が社会学や経済学では良いのか？（対数正規分布の主張も）**
- Long tail**をすくい上げる良い仕組みはあるか？**
  - Amazon.co.jp**も**2ch.net**も**$b < 1$
- **流行とは何か（モデル化），どのように分析するか**

文献

K. Hattori, T. Hattori, Stochastic Processes and their Applications **119** (2009) 966–979.

K. Hattori, T. Hattori, Funkcialaj Ekvacioj **52** (2009) 301–319.

K. Hattori, T. Hattori, RIMS Kokyuroku Bessatsu (2010), to appear.

Y. Hariya, K. Hattori, T. Hattori, Y. Nagahata, Y. Takeshima, T. Kobayashi, Tohoku Mathematical Journal (2011), to appear.

http://web.econ.keio.ac.jp/staff/hattori/amazonj.htm
Google**検索キーワード**　　**服部哲弥**