HOKKAIDO UNIVERSITY

Title	2007年度 混沌系工学特論講義ノート
Author(s)	井上, 純一
Description	http://chaosweb.complex.eng.hokudai.ac.jp/~j_inoue/konton2007/konton2007.html
Issue Date	2007
Doc URL	https://hdl.handle.net/2115/32315
Rights(URL)	https://creativecommons.org/licenses/by-nc-sa/2.1/jp/
Type	learning object
File Information	konton2007_3.pdf, 第3回講義ノート

Hokkaido University Collection of Scholarly and Academic Papers : HUSCAP

混沌系工学特論 講義ノート #3

担当：井上 純一 (情報科学研究科棟 8-13)
URL：http://chaosweb.complex.eng.hokudai.ac.jp/~j_inoue/

平成 20 年 1 月 28 日 (最終回)

目 次

2.4 点過程 23
2.5 ポアソン過程 25
2.6 レート変動間隔の実データ解析 27
2.6.1 ワイブル確率紙の方法 27
2.6.2 変動間隔が揺らぐ GARCH(1,1) 過程での計算機実験 28
2.7 ジニ係数: 格差の一指標 29
2.7.1 解析的評価 30
2.7.2 数値的評価 31
2.8 平均待ち時間 32
2.8.1 レート変動とその観測に関連する時間 33
2.8.2 レートの観測に関するパラドクス 34

2.4 点過程

前回までの講義では株価や為替レート等の金融商品の変動を確率過程にみたてて解析するための基本的な確率モデルを見てきた. このとき, 価格変動が生じる「時刻」とそれらの「間隔」に関しては特に何も言及しなかった. インターネット上で公開されている株式などはその価格が 1 時間程度の間隔で提示されるものもあれば, その履歴が 1 日単位, 1 週間単位などで提示されるものもある. 一方, 円/ドルなどの為替レートでは変動が逐次提示されることはなく, たいていの場合, 1 日を通して同じレートが用いられている[1]. 従って, 顧客に対し, そのように一定の単位時間に区切って価格の変動を提示する限り, その確率モデルを考える際には価格変動がある単位時間 (ステップ) 毎に発生するという仮定を置けば十分であった. 例えば, ウイナー過程は

$$X_{t+1} \quad = \quad X_t + Y_t, \quad P(Y_t) = \mathcal{N}(0, \sigma) \tag{65}$$

で与えられたが, この際, レート X_t の変動は $t+1-t=1$ 単位時間毎に生じる.

しかし, インターネットを介した金融取引きが一般にも普及し, 所謂, インターネット・トレーディング・システムを構築する際には顧客に対して可能な限り頻繁にレートを提示することが重

[1] 例えば, 外国旅行へ行く前に円を外貨に換金する際, 銀行では「本日のレート」というような掲示があるし, 多くの場合, このレートは 1 日中変わらない

要になる．つまり, 極端な場合には市場のレートをそのまま開示する方が, 顧客各自が購入した商品の価格動向を逐次観察することができ, その結果, オンラインでの投資を楽しみたいという類の顧客ニーズに応えることができる．しかし, 現実には, インターネットなどを介して我々が目にすることのできる「開示レート」の背後には (プロのトレーダ達の取り引きによって決まる) 「市場レート」があり, トレーディング・システムを提供する側が, その市場レートからある程度の情報を「間引いて」からそのレートを提示しているわけである．

このとき, どのように情報を間引くかが問題になるが, 一番単純には, 前述のように 1 日単位, あるいは 1 時間単位と一定の時間を区切り, その時点での市場レートを公開するのが良さそうである．しかし, 例えば大きな事件があったり, 何らかのアクシデントでレートそのものが乱高下し, その乱高下がある日の正午から午後 6 時の間継続した場合, 1 日ごとのレート提示 (開示) では, この時間帯のレート乱高下は顧客には見えなくなってしまう．顧客にとっては出来る限り市場レートそのものにアクセスできた方が, このような乱高下のタイミングを利用して (それは小さなチャンスかもしれないが) 自分の資金を増やすことができるかもしれない．しかし, これは銀行 (システムを構築する側) にとってみれば逆な意味でリスクを負うことになりかねない．従って, 両者にとって好ましいレート提示の方法は市場レートそのものを全て提示するのではなく, やはり, 情報を間引くのであるが, その**情報の縮約によって作られる開示レートの変動が市場レートの性質をある程度反映している**ようなものが望ましい．この意味では 1 日毎, 1 時間毎などの一定間隔での情報公開は好ましくない．頻繁に変動したレート変動は上述のように「粗視化」されてしまうからである．このような

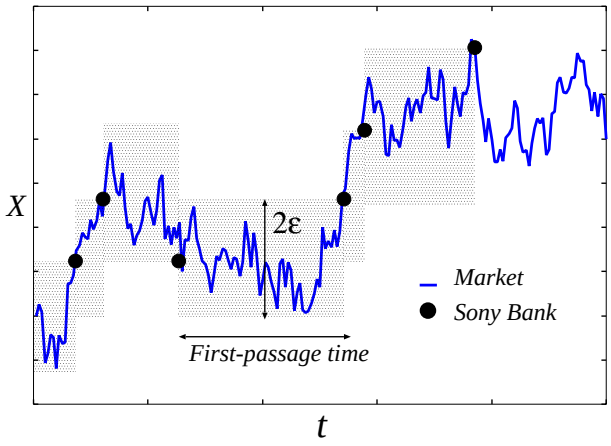

図 14: 市場レート (実線) に対する第一通過過程 (First-Passage Process) を取ったのちの提示レート (黒丸)．このような提示システムを採用している企業としてソニー銀行がある．

市場レートの変動をうまく反映させた情報提示 (開示) として最近, **第 1 通過過程 (First-Passage Process)** を用いる方法が 24 時間随時トレーディングが可能な個人顧客向けのインターネット・バンクであるソニー銀行[2] などで採用されている．これは図のように頻繁に変動する市場レート (実線) に対し, 2ϵ の幅を設定し[3], 市場レートがこの幅を超えた場合のみ, その時点でのレートを顧客に提示する．このやり方だと, ± 0.1 円を越えて大きく変動した後のレートは必ず顧客に提示されることになり, 1 日毎, 1 時間毎にレートを提示するやり方よりも市場レートを開示レートに反映させ

[2] http://moneykit.net
[3] ソニー銀行の円ドル為替レートの場合には $\epsilon = 0.1$ 円と設定されている．

ることができる[4] ここで注意しなければならないのは, 市場レートがウイナー過程 (65) に従っていたとしても, この第一通過過程のフィルタ (以下では単に「フィルタ」と呼ぶ) を介して提示されるレート変動の時間間隔はもはや一定ではなくなっているという事実である. つまり, フィルタを介した後のレート変動が時刻 $t_1, t_2, t_3, \cdots, t_N$ (当然, この順序は $0 < t_1 < t_2 < \cdots < t_N$ であり, $N \to \infty$ とする) に生じたとし, その間隔を

$$\Delta t_i \quad = \quad t_{i+1} - t_i, \quad (i = 1, \cdots, N) \tag{66}$$

で定義すると, Δt_i は一定ではなく, 揺らいでいて, ある確率分布を持つことになる. このレート変動時刻を表す点列: $t_1, \cdots, t_N$ からなる確率過程を**点過程 (Point Process)** と呼ぶ. このような点過程としてスーパーのレジに到達する顧客の時刻, 高速道路の料金所に入ってくる車の到着時刻などが挙げられるが, この他にも我々の身の周りで多くの例を見ることができる.

課題 2 :
身の周りで点過程であると思われる事例を数件具体的に挙げよ. ただし, 上記 2 例を除く.

2.5 ポアソン過程

ここで, レートの変動 (イベント) がランダムに生じる場合にそのレート変動間隔の従う分布を求めてみよう. $P_n(t)$ をそのようなイベントが時刻 0 から t までの間に n 回生じる確率とする. また, 時刻 t と $t + \Delta t$ の間の Δt の時間間隔にそのようなイベントはその間隔に比例するものとし, その比例係数を λ と置く. つまり, t と $t + \Delta t$ の間の微小時間 Δt にレート変動が生じる確率は $\lambda \Delta t$ となる. このとき, 確率の保存から次の方程式が成り立つ.

$$P_n(t + \Delta t) \quad = \quad (1 - \lambda \Delta t) P_n(t) + \lambda \Delta t P_{n-1}(t) \tag{67}$$

この右辺第 1 項は時刻 t までに n 回のイベントが生じている場合, 次の Δt の間隔でイベントが生じなければ左辺への寄与となることから来る部分. 右辺第 2 項は時刻 t までに $n-1$ 回のイベントが生じている場合, 次の Δt の間隔でイベントが生じれば左辺への寄与となることから来る部分であり, この 2 つの寄与の合算が $P_n(t + \Delta t)$ を与えることは明らかであろう. また, 右辺の各項がこのよう書けることから, ここでは「時刻 t と $t + \Delta t$ ではイベントが互いに独立に起こる」という仮定を置いていることに注意しておこう. これを $\Delta t \to 0$ のもとに書き換えると次の微分方程式が得られる.

$$\frac{dP_n(t)}{dt} \quad = \quad -\lambda (P_n(t) - P_{n-1}(t)) \tag{68}$$

これは今までにも度々出てきた**マスター方程式**と呼ばれるものである. 通常, (統計物理や確率的情報処理で扱うような) マスター方程式はその変数が膨大となるため解くことが困難である場合が多

[4] これ以外にもいろいろやり方はあるかもしれない. 例えば, 立て続けに 5 回連続してレートが上昇あるいは下降した場合にはその 5 回目の変動直後のレートを提示するとか, もう少し複雑にしたければ, 1 ステップでのレート変動を「上昇 (u)」「留まり (s)」「下降 (d)」と 3 値に「符号化」し, (u, u, d, u, d, s) の組み合わせが時系列中に現れたら, その時点の値を開示するというような方法もとれるかもしれない. しかし, そのような時系列の構造を反映した複雑なレート開示を採用すれば, どのような状況が生じるのかは直観的にちょっとわかりにくい. その意味で, ここでの第 1 通過過程を用いた情報開示はかなり妥当な方法であると思われる. また, 同じ第 1 通過過程を用いるにしても, 顧客の生活リズム (顧客が日本人だけであると考えれば, 日本時間の真夜中にトレーディングする人はあまり居ない) や実社会での「季節性」を考えて, レート窓の幅 ϵ に時間依存性を持たせるやり方もあるかもしれない.

いが, 上記の場合には解析的にその解を求めることができる. $n=-1$ のとき, $P_{-1}=0$ であるとすると, $n=0$ に対する方程式は

$$\frac{dP_0(t)}{dt} = -\lambda P_0(t) \tag{69}$$

であり, $P_0(0)=1$ の条件下でこの解は $P_0(t)=\mathrm{e}^{-\lambda t}$ である. 従って, $n=1$ のときに解くべき方程式はこの結果を (68) 式に代入して

$$\frac{dP_1(t)}{dt} = -\lambda P_1(t) + \lambda \mathrm{e}^{-\lambda t} \tag{70}$$

が得られる. この方程式は通常よくやるように, 右辺第 2 項を無視した斉次方程式の解の積分定数を時変なものとした解 $A(t)\,\mathrm{e}^{-\lambda t}$ を (70) 式に代入し, A についての微分方程式を作り, それを解けばよい (定数変化法). 実際, その方程式は

$$\frac{dA}{dt} = \lambda \tag{71}$$

となり, その解は $A(t)=\lambda t$. 従って, 求めるべき方程式の解は $P_1(t)=\lambda t\mathrm{e}^{-\lambda t}$ となる. この操作を繰り返すと, $n=k$ のときの方程式の解 $P_k(t)$ は

$$P_k(t) = \frac{(\lambda t)^k}{k!}\,\mathrm{e}^{-\lambda t} \tag{72}$$

となる. この分布を**ポアソン分布**と呼ぶ. また, 上記の意味でイベントがランダムに次々と発生す

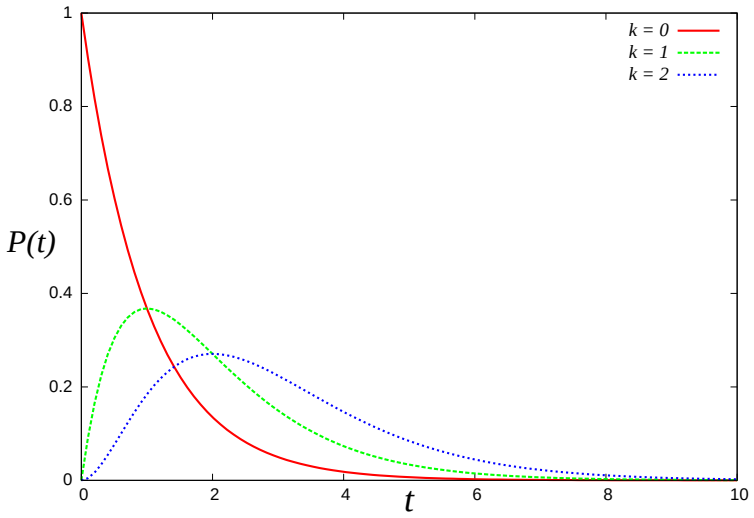

図 15: $\lambda=1$, $k=0,1,2$ の場合のポアソン分布.

る確率過程 (点過程) を**ポアソン過程**という. ここで, このポアソン過程に対し, 計測を始めてからはじめのイベントが生じるまでの時間間隔 (duration) t の従う分布は (72) 式で $n=0$ とおいた

$$P(t) \propto \mathrm{e}^{-\lambda t} \tag{73}$$

のような指数分布に従うことがわかる. 従って, 為替レートの変動間隔 t の従う分布を実データから評価し, それが指数分布に従うのであれば, レートの変動は完全にランダムなポアソン過程であることがわかる.

2.6　レート変動間隔の実データ解析

実際にレート変動がランダムに起こるポアソン過程に従っているのか, あるいはランダムではなく, レート変動間隔間には何らかの相関があるのかを調べるためにはレートの変動間隔の従う分布を実データから評価すればよい. その際, 問題の焦点がレート変動がポアソンか否かであるのであれば, 指数分布を変数 m を用いて一般化した次のワイブル分布:

$$P_{m,a}(t) \quad = \quad \frac{mt^{m-1}}{a}\exp\left(-\frac{t^m}{a}\right) \tag{74}$$

に対し, この分布と実データから得られるレート変動間隔のヒストグラムとの当てはまり具合を調べることでパラメータ m, a を求めればよい (a は分布のスケールを決定するパラメータ. 図 16 参照). ここで, ワイブル分布は $m = 1$ であれば指数分布になり, 従って, その変動間隔分布を持つ確

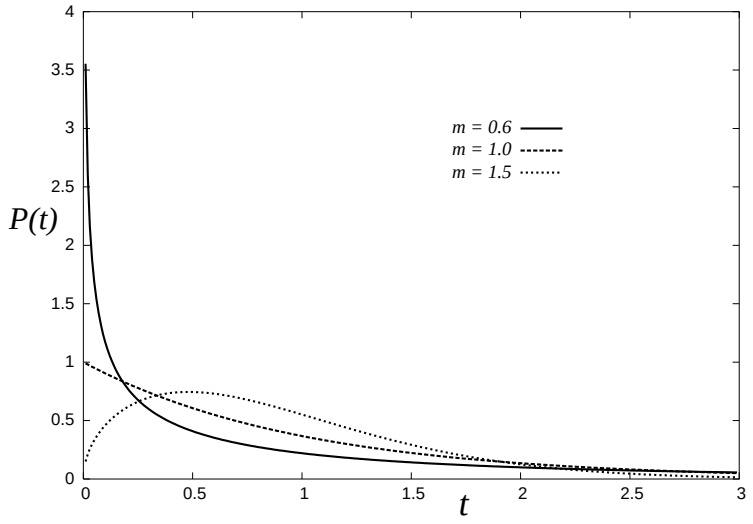

図 16: ワイブル分布.

率過程はポアソン過程ということになる. つまり, パラメータ m の値の $m = 1$ からの「ズレ」によってポアソン過程からの外れ具合を検出しょうというわけである. 「分布間の近さ」を議論するわけであるから, 実データからの分布 (経験分布) と上記ワイブル分布とのカルバック-ライブラ距離を変数 m, a の関数とし, それを最小化するように m, a を決める方法をとることもできるが, ここではもう少し直接的な方法 — ワイブル確率紙の方法 — を紹介することにしょう.

2.6.1　ワイブル確率紙の方法

通常, データ (x, y) が指数分布に従うかどうかを調べる際には $y = \mathrm{e}^{-x}$ の両辺の対数を取った $\log y = -x$ から, このデータを y 軸が対数スケールとなるようにプロットし, そのプロットが直線に乗るかどうかで判断した. 同じような方法がワイブル分布で使えないかを考えたい. このとき, まず, ワイブル分布 (74) 式の累積分布:

$$W(t) \quad = \quad \int_0^t P_{m,a}(t)dt = \exp\left(-\frac{t^m}{a}\right) \tag{75}$$

を計算する. 従って, 次の式が成り立つことが容易にわかる.

$$\log\left(-\log W(t)\right) \quad = \quad m\log t + a \tag{76}$$

よって, レート変動間隔のヒストグラムから累積分布を求め, 間隔 t の「頻度」の対数を取ったものの符号を変えてさらにその対数をとり, それと間隔 t の対数をとったものをそれぞれ y, x 軸にプロットしたとき, そのプロットが直線に乗ればレート変動間隔はワイブル分布に従うことがわかり, このときの直線の傾きがワイブル分布を特徴つけるパラメータ m を決めることになる (y 軸との切片が a を与える). その m が 1 と異なれば, その確率過程はポアソン過程ではないと結論されることになる. この手の方法を**ワイブル確率紙の方法 (Weibull Paper Analysis)** と呼ぶ.

2.6.2　変動間隔が揺らぐ GARCH(1,1) 過程での計算機実験

ところで, 前述のレート変動幅 2ϵ のフィルタの効果はどのようなものであろうか？ 我々はここでレート変動間隔そのものに興味があるので, 確率過程のレート変動間隔がフィルタを介す前後でどのように変化するのかに興味がある. そこで, 前回学んだ GARCH(1,1) 過程の時間間隔がワイブル分布に従う確率過程に対し, 実際にレート窓のフィルタをかけることでこれを調べてみよう. 具体的には次の確率過程にフィルタをかける.

$$X_{t+\Delta t} \quad = \quad X_t + Y_t, \;\; P(Y_t) = \mathcal{N}(0, \sigma_t) \tag{77}$$

$$\sigma^2_{t+\Delta t} \quad = \quad \alpha_0 + \alpha_1 X_t^2 + \beta_1 \sigma_t^2 \tag{78}$$

$$P_{m_0,a}(\Delta t) \quad = \quad \frac{m_0(\Delta t)^{m_0-1}}{a}\exp\left(-\frac{(\Delta t)^{m_0}}{a}\right) \tag{79}$$

ここで, 為替レートそのものと時間間隔の間には何ら相関がないものと仮定する. すなわち, $P(X_t, \Delta t) = P(X_t)P_{m,a}(\Delta t)$ である. しかし, 当然, この相関を取り込んでモデル化することも可能である[5].

さて, この確率過程に窓をかけ, 出てくる確率過程の時間間隔をワイブル確率紙で調べよう. その際, m_0 で指定される第 1 通過過程の背後にある確率過程が, 第 1 通過過程を経ることでその時間間隔の従う統計性に変化が生じ, m_0 とは異なるパラメータ m で指定されるワイブル分布に従うようになると仮定する. そして, 最終的に m_0-m プロットを描くことでパラメータ m の変化を介してここでのフィルタの効果を調べようというわけである. その結果の一例を図 17 に載せる. ここでは $(\alpha_0, \alpha_1, \beta_1) = (0.4, 0.3, 0.3)$ と変数を選び, $m_0 = 0.6, 1.2$ のワイブル分布で変動間隔が揺らぐ GARCH(1,1) 過程のフィルタ出力の変動間隔をワイブル確率紙により評価している. ただし, レート自体は完全な連続値を取るわけではなく, レート窓幅 $\epsilon = \Delta = 0.1$ に応じた離散値を取ることから,

$$X_t \quad = \quad \Psi_\Delta(X_t) \equiv \Delta\,\mathrm{ceil}(\Delta^{-1}X_t) \tag{80}$$

なる変換を施していることに注意しょう. この図より, ワイブル確率紙による解析で出力データは直線上に乗っている. 従って, フィルタにかけた後の確率過程のレート変動間隔もやはりワイブル分布に従う. しかし, その傾きは一般に $m \neq m_0$ である. そこで, 図 18 に m_0-m プロットを載せよう. この図より, フィルタへの入力データが $m_0 = 1$ のポアソン過程である場合でも, フィルタをか

[5] 相関があれば $P(X_t, \Delta t) = P(X_t|\Delta t)P_{m,a}(\Delta t)$ なので, 条件つき確率 $P(X_t|\Delta t)$ を具体的に与えてモデル化することになる. 実際にそうなっているか否かは検証が必要であるが, 例えば, 「長く待たされた直後のレートは大きく変動する」などを具体的に条件つき確率 $P(X_t|\Delta t)$ に反映させることができるかもしれない.

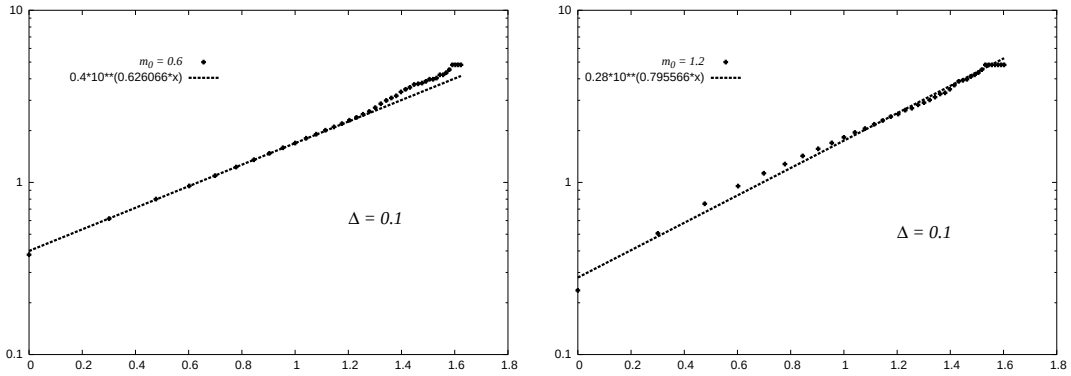

図 17: $m_0 = 0.6, 1.2$ のワイブル分布で変動間隔が揺らぐ GARCH(1,1) 過程のフィルタ出力の変動間隔をワイブル確率紙により評価する.

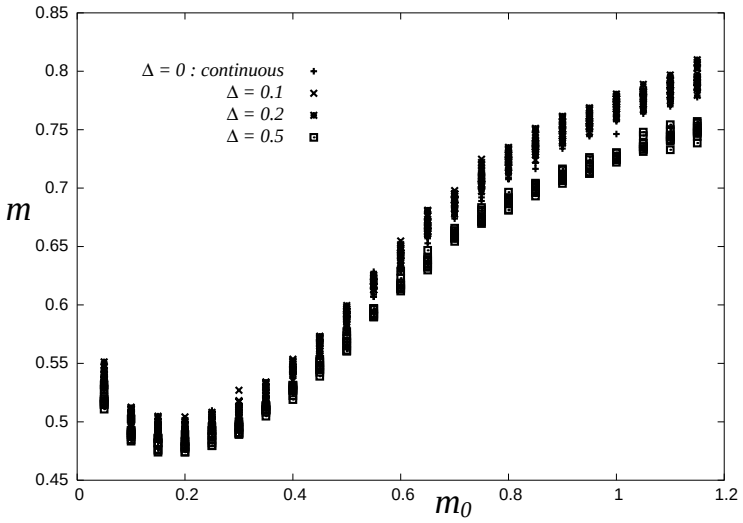

図 18: m_0-m ダイアグラム.

けた後では $m \neq 0$ であり, ランダムなレート変動間隔ではなくなることがわかる. では, レート変動間隔がランダムでなく, ある種の偏りがある場合, それをどのように定量化すればよいのであろうか. 時節ではそれを定量化する一つの指標としてジニ係数について見ていくことにする.

2.7　ジニ係数: 格差の一指標

最近では社会における格差が問題になっており, 各国での格差の指標として**ジニ係数 (Gini index)** と呼ばれる量が用いられている. この量は $[0, 1]$ 間の実数値を取るように規格化されており, 完全に平等な社会で, その意味で富が国民に等分配される (少なくとも社会主義が考える) 理想的社会では 0 をとり, その対極の一人が全ての富を独占する状況では 1 をとる. 従って, ここで言うところの「格差」という観点で捉えるのであれば, ジニ係数がゼロに近ければ近いほど格差の無い社会ということになる. 参考までに 2000 年の統計によると日本のジニ係数は 0.314, アメリカ合衆国は 0.357, メキシコが 0.450, デンマークが 0.225 であった.

2.7.1 解析的評価

ジニ係数を解析的に求めるためには, まず, 所得の分布から**ローレンツ曲線 (Lorentz Curve)** を求めなければならない. そのために所得分布を $P(t)$ とし, その累積分布:

$$X(r) = \int_0^r P(t)dt \tag{81}$$

と規格化された累積所得:

$$Y(t) = \frac{\int_0^r tP(t)dt}{\int_0^\infty tP(t)dt} \tag{82}$$

を計算し, r を媒介変数として X と Y の関係をプロットする. このとき描かれる曲線がローレンツ曲線である. これは $[0,1]$ 間で下に凸の曲線であり, 完全平等の場合には $P(t) = \delta(t - t_*)$ を代入し

$$X(r) = \int_0^r \delta(t - t_*)dt = \Theta(r - t_*) \tag{83}$$

$$Y(t) = \frac{\int_0^r t\delta(t - t_*)dt}{\int_0^\infty t\delta(t - t_*)dt} = \frac{t_*\Theta(r - t_*)}{t_*} = \Theta(r - t_*) \tag{84}$$

となるから, $Y = X$ がローレンツ曲線となる ($\Theta(\cdots)$ は階段関数). これを「完全平等線」と呼ぶことにする.

ワイブル分布 (74) のローレンツ曲線は簡単な計算の結果

$$Y = Q\left(\frac{1}{m} + 1, -\log(1 - X)\right) \tag{85}$$

となる. ここに, $Q(z, x)$ は不完全ガンマ関数であり

$$Q(z, x) = \int_0^x t^{z-1}\,\mathrm{e}^{-t}dt \tag{86}$$

で定義される. 図 19 にいくつかの m 対するワイブル分布のローレンツ曲線を載せる. ジニ係数はこのローレンツ曲線と完全平等線との間の面積を 2 倍したものとして算出される. 従って, ローレンツ曲線自体が完全平等線の場合にはゼロ. その対極の一人勝ちの場合には一人に全ての富 r があつまるから (83)(84) の結果に $t_* = r = \infty$ を代入して, $Y = X = 1$ がローレンツ曲線となり, このときの面積は $Y = X$ と X 軸, Y 軸とで囲まれる三角形の面積 $1/2$ であり, 従ってジニ係数はその 2 倍の 1 となる.

このジニ係数は具体的に書くと

$$G = 2\int_0^1 (X - Y)dX = 2\int_0^\infty (X(r) - Y(r))\frac{dX}{dr} \cdot dr \tag{87}$$

となる. 具体的にワイブル分布に対して計算すると

$$G = 1 - \left(\frac{1}{2}\right)^m \tag{88}$$

が得られる. ここで変数 a には依存しないことに注意しておこう. これを m の関数としてプロットすると図 20 のようになる. ところで, ここでは所得の分布を取り上げたが, 不平等の程度を測る量は所得に限る必要はない. 事実, 例えば, ここで取り上げた所得を「レート変動間隔」と読み直すことで, レート変動の間隔がどの程度ばらつき, どの程度偏りがあるのかを調べることができる. 例

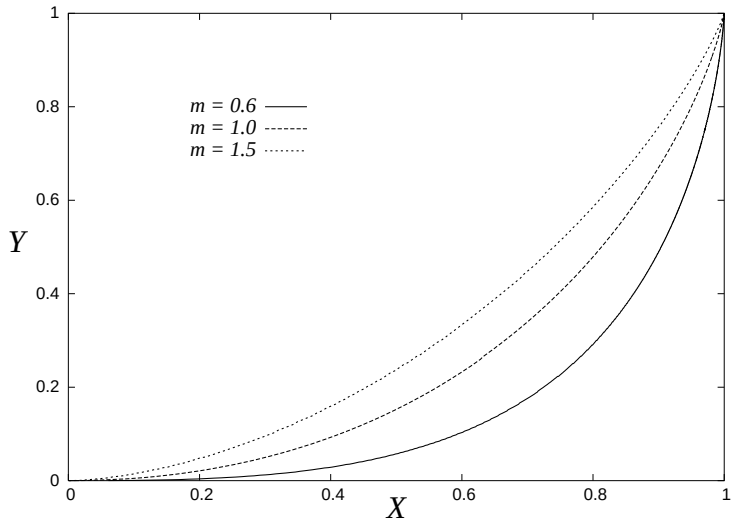

図 19: ワイブル分布に対するローレンツ曲線.

えば, $m = 1$ としてレートの変動がランダムなポアソンに従う場合, ジニ係数は 0.5 となる. また, ジニ係数がゼロに近ければ, おおよそ同じ変動間隔でレートが変化すると言える. 逆にジニ係数が 0.5 より大きければ間隔は偏りが大きく, 飛びぬけて大きな間隔, あるいは逆に小さな間隔がところどころで現れるようになる. ちなみに, 第 1 通過過程を採用しているソニー銀行のレート変動間隔はワイブル確率紙からの結果 $m = 0.59$ より, $G = 0.694$ 程度となり, これは後者, つまり, 飛びぬけて大きな間隔, あるいは逆に小さな間隔がところどころで現れている.

2.7.2 数値的評価

データが与えられたときに数値的にジニ係数を求めることもできる. 例えば各人の所得が $x_1 \leq x_2 \leq \cdots \leq x_N$ のように与えられたとすると (必ずしもデータとしてはこの順序に並んでいないかもしれないが, その場合には逐次ソーティングを行い, この順番に並び替える),

$$X_i = \frac{i}{N} \tag{89}$$

$$Y_i = \frac{\sum_{r=1}^{i} x_r}{\sum_{r=1}^{N} x_r} = \frac{\sum_{r=1}^{i} x_r}{N\mu}, \quad \mu \equiv \frac{1}{N}\sum_{r=0}^{N} x_r \tag{90}$$

であるから, $dX_i = (i+1)/N - i/N = 1/N$ に注意して

$$G = 2\sum_{i=1}^{N}(X_i - Y_i)\frac{1}{N} = \frac{1}{N^2\mu}\sum_{i=1}^{N}(2i - N - 1)x_i \tag{91}$$

が求めるジニ係数となる.

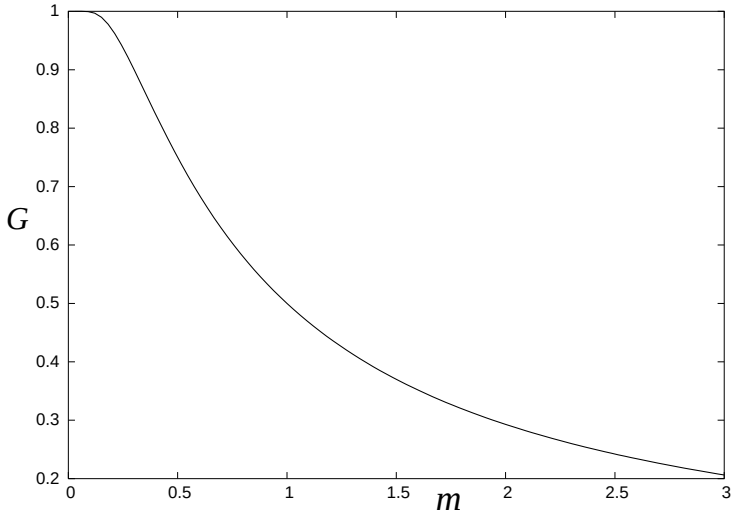

図 20: ワイブル分布に対するジニ係数.

課題 3 :
混沌系工学特論のページ:

`http://chaosweb.complex.eng.hokudai.ac.jp/~j_inoue/konton2007/konton2007.html`

上に, ある分布から計算機を用いて人工的に生成したデータ (kadai3DATA.txt) を載せるので, このデータのジニ係数を (91) 式を用いて求めよ.

2.8 平均待ち時間

前節までで, 確率過程としてみたときの為替レートの変動間隔は一定ではなく, 揺らぎを持っていることがわかった. また, その揺らぎをジニ係数を介して定量化し, インターネット・トレーディングシステムを採用しているソニー銀行のそれは 0.694 であることも見た. この結果からレート変動がランダムに起こるポアソン過程とは異なり, 飛びぬけて大きな間隔, あるいは逆に小さな間隔がところどころで現れていることになる. ところで, ワイブル確率紙によるデータ解析でソニー銀行のレート変動間隔は $m = 0.59, a = 50.858$ のワイブル分布で特徴つけられることがわかったので, この分布の平均:

$$\langle t \rangle \quad = \quad \int_0^\infty t P_{m,a}(t) dt \tag{92}$$

を計算することで平均レート変動間隔を評価することができる. ソニー銀行の場合にはこれが理論, 実データ解析双方から約 20 分となることがわかっている. この量はシステム設計においては有益な情報に違いないが, 顧客にとってさえも, この情報が有益であるかどうかはわからない[6]. 例えば, 顧客は任意の時刻にインターネットにアクセスし, 為替レートをチェックするわけであるが, もし, 仮に変動間隔に偏りを持つソニー銀行のレートをチェックしたとすると, 極端に変動間隔の長いタ

[6] 有益には違いないだろうが, 他により適切な量があるのではないかと, ここでは考える.

イミングでレートをチェックしてしまい, 長い待ち時間にうんざりしてしまうかもしれないし, 逆に, 極端にレート変動間隔が短いタイミングでレートをチェックし, あまりにも早いレートの変化に驚くことになるかもしれない. 顧客のレート・チェックが 1 日 24 時間の中でどの時間帯にも偏ることなく行われる条件下では, レート変動間隔がランダムならば, 前述の平均変動間隔が顧客の平均待ち時間と考えても良いであろう. しかし, レート変動がポアソン過程ではなく変動間隔に偏りがある場合には顧客の平均待ち時間は平均変動間隔とは異なる値を持つかもしれない.

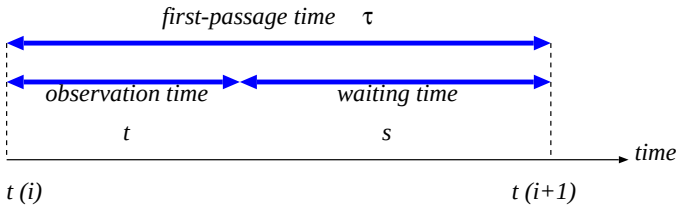

図 21: レート変動とその観測にともなう 3 つの時間.

2.8.1 レート変動とその観測に関連する時間

そこで, もう一度問題を整理すると, 我々の扱っている系にはレートの変動とその観測に関連して 3 種類の異なる時間 (時刻) が存在することがわかる. 一つは第一通過時間 τ であり, この時間の従う分布は実データ解析からワイブル分布であることがわかっている. もう一つは, 直前のレート変動時刻から測って顧客がレートをチェックするまでの時間 t であり, 最後は, レートをチェックしてから実際にレートが変動するまでの時間 s である. これら 3 者の関係は図 21 で与えられる.

ここで, 我々はインターネット・トレーディングを行う顧客にとって重要な情報は τ よりも s であると考える. つまり,

$$s \quad = \quad \tau - t \tag{93}$$

の分布, あるいは, この平均値こそ重要であると考えるわけである.

そこで, t の分布を $P_O(t)$ とし, 問題の待ち時間 s の分布 $\Omega(s)$ をこれと $P_{m,a}(\tau)$ を用いて表すと

$$\Omega(s) \quad \propto \quad \int_0^\infty d\tau \int_0^\tau dt\, Q(s|\tau,t) P_O(t) P_{a,m}(\tau). \tag{94}$$

となる. ここに, $Q(s|\tau,t)$ は第一通過時間が τ である間隔に直前のレート変動から t だけ経過した時点で顧客がログインし, レートをチェックした条件下でその顧客が次のレート変動まで s だけ待つ確率であり, ここでは

$$Q(s|\tau,t) \quad = \quad \delta(s-\tau+t) \tag{95}$$

で与えられる. 従って, 待ち時間 s の分布は規格化定数も含めて

$$\Omega(s) = \frac{\int_0^\infty d\tau P_{m,a}(\tau) \int_0^\tau dt\, \delta(s-\tau+t) P_O(t)}{\int_0^\infty ds \int_0^\infty d\tau P_{m,a}(\tau) \int_0^\tau dt\, \delta(s-\tau+t) P_O(t)} \tag{96}$$

と書くことができる. 顧客のログイン時間は一様: $P_O(t) = 1$ と考えるのが自然であろうから, これを考慮して問題の待ち時間 s の平均を計算してみると

$$w \quad = \quad \langle s \rangle = \int_0^\infty ds s \Omega(s) = \frac{\int_0^\infty ds s \int_s^\infty d\tau P_{m,a}(\tau)}{\int_0^\infty ds \int_s^\infty d\tau P_{m,a}(\tau)}$$

$$= \frac{\int_0^\infty \frac{d}{dt}\{s^2/2\}ds \int_s^\infty d\tau P_{m,a}(\tau)}{\int_0^\infty \frac{d}{ds}\{s\}ds \int_s^\infty d\tau P_{m,a}(\tau)} = \frac{(1/2)\int_0^\infty s^2 P_{m,a}(s)ds}{\int_0^\infty s P_{m,a}(s)ds} = \frac{E(\tau^2)}{2E(\tau)} \tag{97}$$

が得られる[7]. ここに第 1 通過時間の分布でのモーメントを

$$E(\tau^n) = \int_0^\infty ds s^n P_{m,a}(s) \tag{98}$$

で定義したことに注意しよう. 従って, (97) 式から顧客の平均待ち時間はここでの第 1 通過時間の従う分布の 2 次と 1 次のモーメントの比で与えられることになる. また, この結果から一般には $E(\tau) \neq w$ であることもわかる.

そこで, 実際に第 1 通過時間の分布をワイブル分布にした場合の平均待ち時間を求めてみよう. 簡単な計算の結果

$$\Omega(s) = \frac{m\,\mathrm{e}^{-s^m/a}}{a^{1/m}\Gamma\left(\frac{1}{m}\right)} \tag{99}$$

$$w = a^{1/m}\frac{\Gamma\left(\frac{2}{m}\right)}{\Gamma\left(\frac{1}{m}\right)} \tag{100}$$

$$\sigma = \frac{a^{1/m}\sqrt{\Gamma(1/m)\Gamma(3/m) - \Gamma(2/m)^2}}{\Gamma(1/m)} \tag{101}$$

を得ることができる. ここに, $\Gamma(x)$ は

$$\Gamma(x) = \int_0^\infty dz z^{x-1}\mathrm{e}^{-z} \tag{102}$$

で定義されるガンマ関数であり, σ は待ち時間の標準偏差である. 図 22(左) に $\Omega(s)$ の概形を載せる.

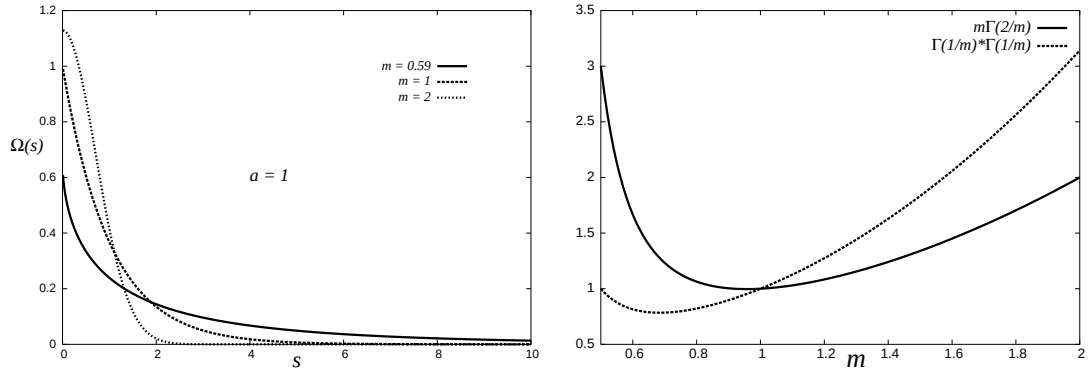

図 22: 待ち時間の分布 $\Omega(s)$ (左). 右図は $E(\tau)$ と w の大小関係. $m = 1$ で両者は一致し, $m < 1$ で $E(\tau) < w$ となる.

2.8.2 レートの観測に関するパラドクス

ところで, 前に一般には $E(\tau) \neq w$ であると述べたが, 特殊なパラメータの選び方に対しては $E(\tau) = w$ となることがあるかもしれない. そこで, $a = 1$ で固定し, m を様々変えた場合に対し,

[7] 待ち時間 w の分布 $\Omega(s)$ までは必要なくて, s の平均値のみを求めるだけでよければ, **講義スライド** #3 の 10 ページに幾何学的証明を載せたので参考にするとよい.

$E(\tau)$ と w の関係を調べてみると図 22 (右) のようになる. つまり, この図では $E(\tau)$ を具体的にワイブル分布に対して計算すると

$$E(\tau) \quad = \quad \int_0^\infty d\tau \tau P_{m,a}(\tau) = a^{1/m} \frac{1}{m} \Gamma\left(\frac{1}{m}\right) \tag{103}$$

となるから, これと w を比較し, $w > E(\tau)$ ならば, $m\Gamma(2/m) > \{\Gamma(1/m)\}^2$ となることから, $m\Gamma(2/m)$ と $\{\Gamma(1/m)\}^2$ の比較でもって, 両者の関係を示したものが図 22 (右) である. この図より, 我々の当初の推測通り, $m = 1$ のポアソン過程の場合には平均待ち時間が平均レート変動間隔と等しくなる. しかし, $m \neq 1$ であれば両者は異なり, 特に $m < 1$ では平均待ち時間が平均レート変動間隔よりも長くなっている. これは一見すると矛盾しているように思える. なぜならば, 図 21 から明らかなうに, 必ず顧客は第 1 通過時間のいずこかへログインするので, s が t よりも大きくなってしまうのは不自然に思えるからである. このことを**レート観測上のパラドクス (Inspection Paradox)** と呼ぶ.

このパラドクスは以下のようにして説明がつく. ジニ係数の m 依存性を示す図 20 より, m の値が 1 から小さくなればなるほどジニ係数が上昇し, その意味でレートの変動間隔に大きな偏りができる. その結果, 極端に長い変動間隔, 極端に短い変動間隔が現れ, 前者のタイミングでログインするチャンスは後者のタイミングでログインするチャンスと比べて圧倒的に多くなることが予想できる. 従って, このようなログインのタイミングも含めて平均すると, 結果としてその待ち時間が変動間隔幅よりも大きくなることがありうるわけである.

ちなみに, w の式にソニー銀行のデータから見積もることのできるパラメータの値 $m = 0.59, a = 50.858$ を代入するとおおよそソニー銀行の顧客の待ち時間は 42 分と評価できる. 平均レート変動間隔は約 20 分だったので, この場合には上記パラドクスが起こる. ところで, この待ち時間 42 分は当然, 背後にある市場データを加工する際に用いたフィルタのレート変動窓の幅 ϵ に依存して変化する. 従って, システム設計の観点からはこの幅 ϵ の制御指針が得られることが望ましい. そのためには背後にある市場レートを調べなければならないが, そのようなデータは今のところ提供されていない. これは全く残念な話ではあるが, 無いものは無いわけであるから, 限られた状況証拠から市場レートを近似する確率モデルを構築していかなければならない. その際に第 1 回の講義でみたマイノリティ・ゲームなどのエージェント・ベースの計算機シミュレーションがそのヒントを与えるかもしれないし, ものすごく楽観的に考えれば, それと実データである為替レートなどをつき合わせることで [ミクロ情報] と [マクロ情報] 間のつながりを明らかにできるかもしれない. これらは今後の課題として残されている.

その後の研究経過に関して興味をもたれた方は時々

`http://chaosweb.complex.eng.hokudai.ac.jp/~j_inoue/major.html`

をのぞいてみてください.

課題 4 :
各自の取り組んでいる研究課題において実データからの知見を使う場合, あるいは, 現在使ってなくとも, 将来, そのような知見を使うことで理解が深まったり理論の正当性が確認できたりすると思われる場合, それらデータの入手方法 (あるいは計測方法), 解析方法等を A4 用紙 2 ページ程度で簡単に説明せよ.

レポート締め切り: 2 月 15 日 (金) 午後 5 時 (厳守). 提出場所は情報科学研究科棟 8-13 のポスト