

解 説 基 礎

Surrogate data 法を用いた生体信号解析

鯉岡 直人¹⁾ 網崎 孝志²⁾

要 旨 生体情報を簡便に把握するため生体信号測定が臨床医学で利用されている。波形の特徴を定性的に利用する以外に、定量的にも、一定のサンプリング時間で生体信号を等間隔の時系列データにして度数分布、平均、分散、標準偏差、中央値などの一般的な統計指標が計算される。また、線形解析法の1つであるフーリエ変換を用いた周波数分布解析(スペクトル解析)が行われる。近年、生体活動は何らかの決定論を背景にもつ非線形的なものが大多数であると考えられ、生体信号の非線形的特徴を調べる解析法が開発された。本稿では測定した時系列データの非線形性を統計学的に検定する surrogate data 法に関して概説する。

鯉岡 直人 網崎 孝志：Surrogate data 法を用いた生体信号解析，呼吸 32(2)：126—135，2013
 キーワード：approximate entropy surrogate data 法 時系列データ 非線形解析

はじめに

生体情報を簡便に把握するため、呼吸運動、心電図、脳波、酸素飽和度など多くの生体信号が臨床医学で利用されている。横軸を時間、縦軸を測定値として時間軸に沿って監視することが多い。生体信号は波形の特徴を定性的に利用するだけでなく、一定のサンプリング時間で等間隔の時系列データ(time series)にして解析される¹⁾。代表的な解析法は時系列データの度数分布、平均、分散、標準偏差、中央値などの統計指標であるが、線形解析法の1つであるフーリエ変換を用いた周波数分布解析(スペクトル解析)も

利用される²⁾。ところで、線形という語句は種々の場面で使用されるが、基本的には、加法性が成り立つことをいう。例えば、システムに対する入力と出力の関係が比例関係になっている場合は、このシステムを線形システムという。逆に、非線形システムでは入力と出力の関係が比例関係にない¹⁾。

近年、生体活動を何らかの決定論を背景にもつ複雑系と仮定し、非線形的特徴を調べる方法が生体信号解析に応用されている。このような解析を非線形解析(nonlinear analysis)というが、この場合の非線形は、生体システムのダイナミクス(時間発展)が線形ではないことを意味していると思われる。複雑系を評価する統計量として生体信号に適用されているのは、相関次元(correlation dimension：D2)³⁾、最大リアプノフ指数(maximum Lyapunov exponent)⁴⁾、approximate entropy (ApEn)⁵⁾、sample entropy⁶⁾などである。

実際には、生体の複雑な時系列データは、その起源が何らかの非線形決定論にあるのか、それとも単に確率的な要因で複雑な様相を呈しているのかを明らかにするため、統計解析のブートストラップ法に似たサロゲートデータ

Analysis of biological signals using surrogate data method

1) 鳥取大学医学部保健学科検査科学技術専攻病態検査学講座
Naoto Burioka

Division of School of Health Science, Department of Pathobiological Science and Technology, Faculty of Medicine, Tottori University, Tottori 683-8503, Japan

2) 鳥取大学医学部保健学科検査科学技術専攻生体制御学講座
Takashi Amisaki

Division of School of Health Science, Department of Biological Regulation, Faculty of Medicine, Tottori University, Tottori 683-8503, Japan

2012年7月23日受付，11月1日採用

(surrogate data)法が開発された⁷⁾⁸⁾。Surrogate data 法を利用して時系列データの解析を新しい観点から行える^{8)~10)}。医学分野においても生体活動に複雑系が関与しているならば、脳波、心電図、呼吸運動など多くの生体信号は生体のダイナミクスから得られた情報の一部分のみであると考えられる。これら断片的な情報から元システムの特徴を推定するために surrogate data 法は役立つかもしれない。本稿では surrogate data の種類と非線形性検定法を概説し、生体信号に適用した例を提示する。

I. Surrogate data の原理

生体信号の変動は周期性をもつようにみえても一定ではなく複雑な挙動を呈する。詳細な検討は、時系列データの平均、分散、度数分布以外に周波数分布解析など線形理論に基づく解析が行われてきた。しかしながら、一般的に生体は非線形システムであり測定された生体信号は非線形現象に起因していると推測される。実際、従来の線形解析では生体信号の複雑性を十分に解析できなかったため、1990年前後から Takens の埋め込み定理¹¹⁾による新しい種類の非線形解析法、即ち、時系列データに時間遅れを導入し高次元ベクトルに再構成して高次元座標系に埋め込むことによりダイナミクスを調べる方法が考案された^{3)~6)}。この方法は決定論的カオス(deterministic chaos)の解析に利用されてきた。カオスとは数学用語で少数自由度を有する決定論的非線形システムから生じる複雑な現象である⁷⁾。一方、不完全に適用して生体信号が何でもカオスであるとの報告も多く認められた⁹⁾。実際、線形確率過程から生じた時系列データも一見複雑な挙動を示す。カオスは必ず非線形性を有するので、複雑系を評価する統計量の妥当性を統計学的に検定する方法が必要とされた。考案された時系列データの非線形性を検定する方法は、測定された生体信号の時系列データに対する線形確率過程の存在を帰無仮説として棄却検定を行う surrogate data (サロゲート データ)法である⁷⁾⁸⁾。即ち、シミュレーションにより、オリジナルデータの線形特性だけを再現するような代理(サロゲート)データを多数発生させ、オリジナルと代理データそれぞれについて複雑系を評価する統計量を計算して仮説の検定を行う^{7)~10)}。

帰無仮説からも分るように、この手法は、与えられた時系列に特定の非線形性が含まれることを証明するためのものではなく、時系列が線形確率過程のものではないことを示すためのものである。なお、線形確率過程とは、時系列の要素が同一正規分布に従う独立な確率変数の線形結合(和)として表すことができるような時系列をいう²⁾¹²⁾。

II. Surrogate data の種類

新しい種類の時系列データの非線形解析は、Takens の埋め込み定理¹¹⁾を用いて高次元の時間遅れ座標系に軌跡を再構成して非線形ダイナミクスの特徴を表す統計量の相関次元、最大リアプノフ指数、ApEn、sample entropyなどを計算する。しかし、生体信号はノイズが多く、計算された統計量の妥当性が常に問題となってきた。特に最大リアプノフ指数はノイズの多い生体信号から計算されることは困難であるが、多くの論文報告があり生体信号は決定論的カオスであるという根拠とされてきた。これらの統計量が非線形決定論をもったシステムから生み出されているならば、少なくとも時系列データが線形確率過程でなく非線形性を保持していることを証明しなければならない。解決策として、測定された生体信号の時系列データに対する線形確率過程の存在を帰無仮説として棄却検定を行う複数の surrogate data 法⁷⁾⁸⁾が考案されている。

時系列データ $u(k)$, ($k=1, 2, 3, \dots, n$) の surrogate data は元の時系列データ(original data), $u(k)$ の統計量を保持している必要がある。統計量として $u(k)$ の平均、分散、自己相関関数、さらには度数分布を考えるとこれらすべてを surrogate data が保持していることが望ましい⁷⁾。特に自己相関関数は Wiener-Khinchine (ウィナー・ヒンチン) の公式により自己相関関数が同じならば周波数分布(パワースペクトル)も同じである²⁾。したがって original data と surrogate data の自己相関関数を同じにするためにはパワースペクトルが同じになるように作成したほうがよい⁸⁾。中心となるのは Fourier transform (FT) surrogate の概念で図 1 のように作成される。以下に代表的な surrogate data を示す(表 1)。

1. Random shuffle (RS) surrogate

時系列データ, $u(k)$, ($k=1, 2, 3, \dots, n$) の各値をランダムに入れ替えたものである。例えると各時刻に測定された値を書き込んだカードをトランプのようにランダムにシャッフルして再度並べたものである。RS surrogate は時間的に無相関なものとなり、重要な自己相関が消失してしまう(図 2)。通常、使用されない。

2. Fourier transform (FT) surrogate

帰無仮説を「測定された生体信号の時系列データは線形確率過程から得られたものである」として想定する場合、surrogate data は線形的特徴を有しながら乱数を作成過程で使用し、original data と同じ自己相関関数をもっている必要がある⁷⁾⁸⁾。そのようなデータは、図 1 に示したような過程を経て得られる。全く同じ周波数分布をもった

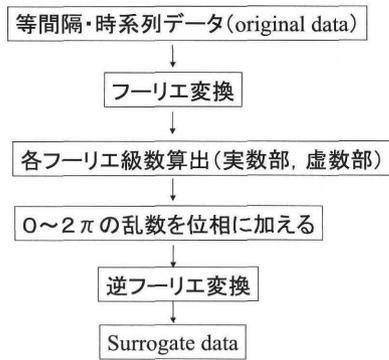


図1 Fourier transform (FT) surrogate の概念

Original data から線形解析のフーリエ変換を行い、位相に乱数を加えて phase shuffle を行う。次いで逆フーリエ変換すると周波数分布が全く同じ surrogate data が得られる。しかし、度数分布は異なってしまう。

め自己相関関数も original data から得られるものと同一である。一方, surrogate data は正規分布に従うため、度数分布は original data とは異なってしまう(図2)。なお、線形確率過程の統計学的モデルとして利用される自己回帰モデルは、自己相関と関係しており、このモデルに基づく検定法も報告されている。

3. Amplitude adjusted Fourier transform (AAFT) surrogate

FT surrogate は original data と同じ自己相関関数をもつが、度数分布は異なってしまう。度数分布を同じにするため、original data を一定の規則でシャッフルする方法が考案された。AAFT surrogate は時系列データ値の大きさの並び(rank order)に着目した手法である⁸⁾。作成法は以下ようになる。① original data のデータ数と同じガウス型乱数(正規分布になる乱数)を作成する、② original data の rank order に従うように①の乱数の並びを入れ替える、③ 並び替えた乱数の FT surrogate を作成する、④ 作成された FT surrogate の rank order に従って original data を並び替える⁷⁾⁸⁾。これらの操作によって平均、分散、度数分布は保存される。しかし、自己相関関数は若干異なってしまう、結果として周波数分布もわずかに異なる(図3)。

4. Iteratively refined surrogate(iterative AAFT)

AAFT surrogate は自己相関関数、周波数分布が original data と若干異なる欠点があった。Shreiber は FT

surrogate と rank order を繰り返し用いて、度数分布が同一で自己相関関数および周波数分布が極めて相似した surrogate data を作成する方法を開発した¹⁰⁾¹³⁾。作成法は以下ようになる。① original data から FT surrogate data を作成するが前述したように度数分布が異なる、② original data を FT surrogate data の rank order に従うように並びを入れ替える。これによって度数分布は original data と同等になるが、今度は周波数分布が若干異なってしまう、③ この並び替えたデータを元に FT surrogate data を再び作成する、④ 作成された FT surrogate data の rank order に従って先程②で作成したデータを並び替える。⑤ これら②~④の操作を10回程度繰り返す⁷⁾¹⁰⁾。これらの操作によって度数分布が同じで周波数分布も極めて相似した surrogate data が作成される¹⁰⁾¹³⁾(図3)。

Ⅲ. Surrogate data を用いた時系列データの非線形性検定

元の時系列データを original data とすると、まず original data から相関次元³⁾、最大リアプノフ指数⁴⁾、ApEn⁵⁾、sample entropy⁶⁾などのうち1つを計算する。しかし、適切な最大リアプノフ指数を生体信号から求めるのは困難であり、相関次元も求まらない例があるため ApEn⁵⁾、sample entropy⁶⁾が適している¹⁴⁾¹⁵⁾。特に ApEn は臨床生分野での報告が多い。統計解析のブートストラップ法に似た surrogate data 法は surrogate data を複数作成できる。帰無仮説を棄却検定するため両側検定を行う場合、正規分布を仮定すると最低39個の surrogate data が必要であると報告されている¹⁰⁾。以下に ApEn を利用して Theiler's sigmas を算出し非線形性検定法に用いた例を解説する¹⁶⁾¹⁷⁾。

1. Approximate entropy

Pincus が複雑系解析を臨床生理学に应用するため ApEn を考案した⁵⁾¹⁸⁾。ApEn は変動が規則正しい時系列データ(予測性がよいもの: ApEn は小さくなる)から変動が乱雑な時系列データ(予測性が悪いもの: ApEn は大きくなる)まで評価できる。Pincus は Korgomorov-Sinai entropy あるいは Eckmann-Ruelle entropy に近似している点で“approximate”を論文中で使用したと思われる⁵⁾。ApEn の計算過程は以下ようになる。

時間間隔が等間隔な時系列データ $u(i)$, ($i=1, 2, 3, \dots, n$) を考える。 $u(i)$ には信号値が入る。等間隔時系列データから Takens の埋め込み定理¹¹⁾を用いた m 次元空間用の再構成ベクトルセット X_i , ($i=1, 2, 3, \dots, n-(m-1)\tau$)は以

表1 各 surrogate data で保存される統計量

種類	平均	分散	度数分布	自己相関関数
RS	○	○	○	×
FT	○	○	×	○
AAFT	○	○	○	△
IR	○	○	○	○

Original data と比較して○：保存, △：ほぼ保存, ×：保存されない
 RS：random shuffle surrogate, FT：Fourier transform surrogate,
 AAFT：amplitude adjusted Fourier transform surrogate, IR：iteratively refined surrogate

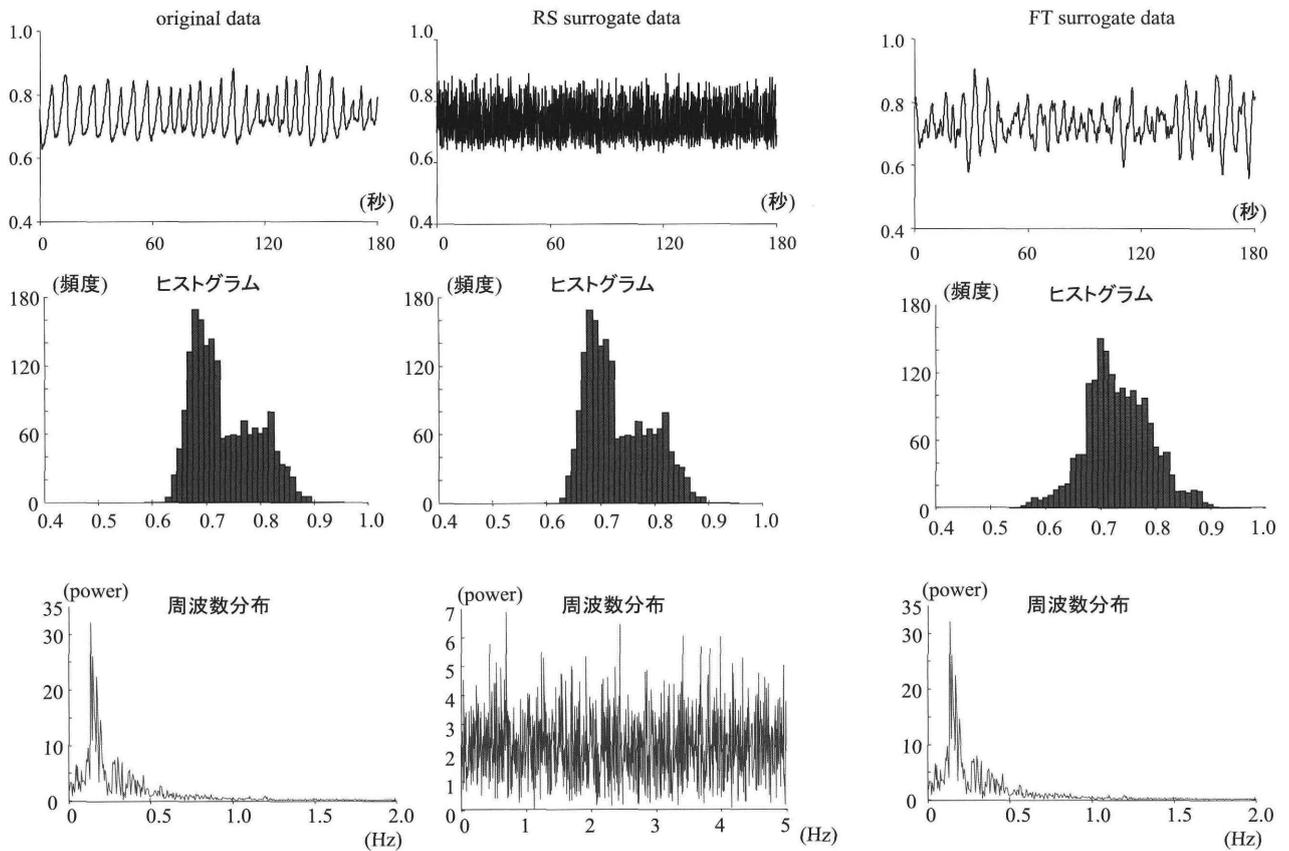


図2 Random shuffle surrogate と Fourier transform surrogate のヒストグラム, 周波数分布

Original data は呼吸運動を 180 秒間, 10 Hz (0.1 秒) で等間隔の時系列データにしたものを使用した(データ数, 180×10=1,800 個)。Random shuffle surrogate は original data の順番を全くランダムに入れ替えるだけなのでヒストグラムは同じである。しかし時間的に無相関になり自己相関が消失してしまうため周波数分布が異なる。Fourier transform (FT) surrogate は周波数分布が全く同じ surrogate data が得られるがヒストグラム(度数分布図)は異なる。

下のようになる。

$$X_i = \{u(i), u(i+\tau), u(i+2\tau), \dots, u(i+(m-1)\tau)\},$$

$$(i=1, 2, 3, \dots, n-(m-1)\tau)$$

時間遅れ座標系の時間遅れ(time lag)の τ は正の整数の定数である。 τ は時系列データの自己相関関数が最初に

$1/e$ に減衰する時間を参考にすることが多い⁷⁾。時間遅れの実際の時間は $\tau \times \Delta t$ となる。 Δt はサンプリング時間である。図4に再構成ベクトルセット作成と埋め込みの概念図を示す。

Pincus は ApEn 考案時, $\tau=1$ の場合を考えた。 $\tau=1$ のとき時系列データ $u(i)$ から作成した再構成ベクトルは,

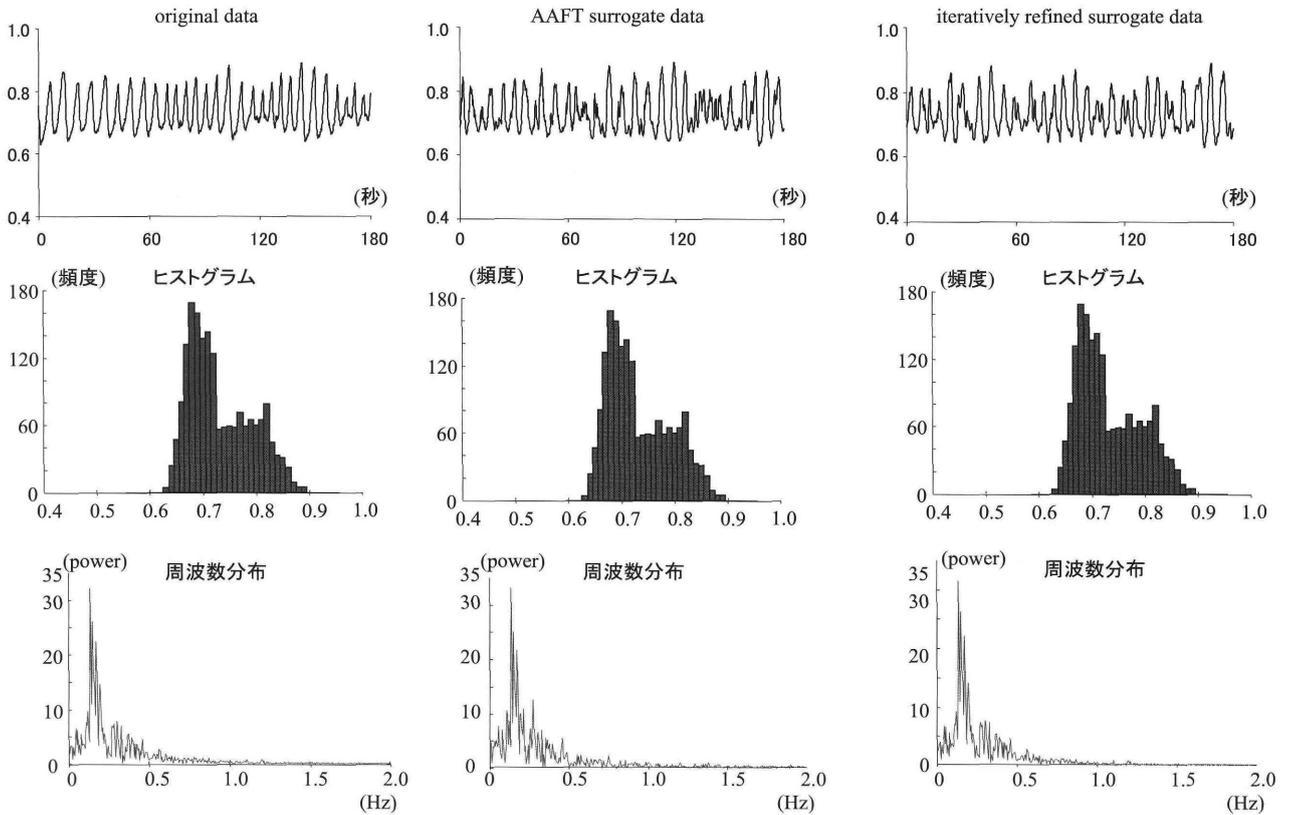


図3 Amplitude adjusted Fourier transform surrogate と iteratively refined surrogate のヒストグラム, 周波数分布
 AAFT surrogate は original data を一定の規則でシャッフルするので surrogate data のヒストグラム(度数分布図)は original data と同一になる。一方, 自己相関関数は若干異なってしまい, 結果として周波数分布もわずかに異なる。Iterative refined surrogate (iterative AAFT) も original data を一定の規則でシャッフルするので surrogate data のヒストグラムは original data と同一になる。また, 自己相関関数が極めて近似するまで繰り返し計算をする。

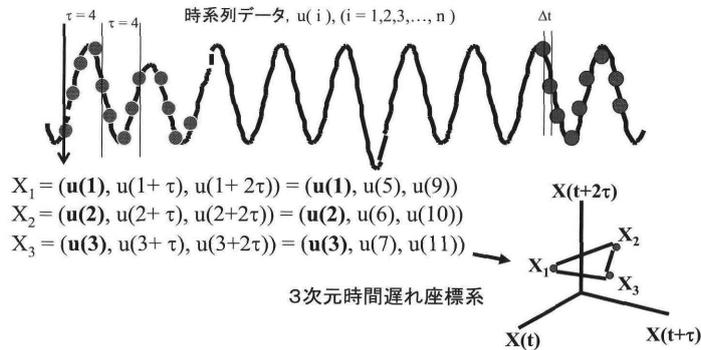


図4 時系列データから再構成ベクトルを作成し時間遅れ座標系に埋め込む概念図
 測定した生体信号を Δt のサンプリング時間で等間隔時系列データ, $u(i), (i = 1, 2, 3, \dots, n)$ にすると隣り合った時系列 $u(k)$ と $u(k + 1)$ の時間差は Δt である。 $u(i)$ には測定された値が入る。図中の時間遅れ τ (time lag) は正の整数である。時間遅れの時間は $\tau \times \Delta t$ となる。例として図では $\tau = 4$ にして, 再構成ベクトルを3次元の時間遅れ座標系に埋め込んだ。3次元の時間遅れ座標系 ($m = 3$) の空間軸は $X(t), X(t + \tau), X(t + 2\tau)$ である。空間軸の時間遅れは時間 ($\tau \times \Delta t$) を使用することもある。

$$X_i = \{u(i), u(i+1), u(i+2), \dots, u(i+(m-1))\},$$

$$(i=1, 2, 3, \dots, n-(m-1))$$

ApEn は以下のように計算される。

$$\Phi^m(r) = (n-(m-1))^{-1} \cdot \sum_{i=1}^{n-(m-1)} \ln C_{m_i}(r),$$

$$\text{ApEn}(n, m, r) = \Phi^m(r) - \Phi^{m+1}(r)$$

$$= (n-(m-1))^{-1} \cdot \sum_{i=1}^{n-(m-1)} \ln C_{m_i}(r) - (n-m)^{-1} \cdot \sum_{i=1}^{n-m} \ln C_{m+1_i}(r)$$

$C_{m_i}(r)$ は以下で表される⁵⁾¹⁸⁾。

まず, $d[X_i, X_j] = \max_{k=1,2,\dots,m} (|u(i+(k-1)) - u(j+(k-1))|)$ を定義する。 $d[X_i, X_j]$ はベクトル X_i の要素, $u(i+(k-1))$ とベクトル X_j の要素, $u(j+(k-1))$ の差の絶対値のうち, 埋め込み次元 m 以下の場合で最大の値を意味する。これを用いて,

$$C_{m_i}(r) = (n-(m-1))^{-1} \cdot \sum_{j=1}^{n-(m-1)} \theta(r-d[X_i, X_j])$$

n は時系列データのデータ数, m は埋め込み次元(embedding dimension), r は定数, \ln は自然対数(natural logarithm)である。 $d[X_i, X_j]$ は, ベクトル X_i とベクトル X_j 間の距離に関連する。 $\theta(t)$ はヘビサイド関数で, $t \geq 0$ ならば $\theta(t)=1$ とし, $t < 0$ ならば $\theta(t)=0$ とする。 $\sum_{j=1}^{n-(m-1)} \theta(r-d[X_i, X_j])$ はベクトル X_i を中心として X_i から一定の定数 r 内に $d[X_i, X_j] \leq r$ を満たす他のベクトル X_j , ($j=1, 2, 3, \dots, n-(m-1)$) が何個含まれていたかを意味している。通常, $m=2, r=0.1 \sim 0.25 \times \text{SD}$ (SD は対象とした時系列データの標準偏差)の定数にして計算する⁵⁾¹⁸⁾。

定義式の意味として再構成ベクトルを m 次元の時間遅れ座標系の中に描写された点とすると, ある点(ベクトル X_i)を中心として, 一定の距離 r 内にどれくらい他の点, ベクトル X_j ($j=1, 2, 3, \dots, n-(m-1)$) が含まれているかを確率 $C_{m_i}(r)$ として計算し, $\ln C_{m_i}(r)$ とエントロピーの形にしている (\ln : 自然対数)。さらに, m 次元の時間遅れ座標系に埋め込まれた点すべてに対して同様の操作を行い, 平均を計算した統計量が $\Phi^m(r)$ である。エントロピーの符号を正にするためマイナス符号をつける ($-\Phi^m(r)$)。そして, 時間遅れ座標系の埋め込み次元 m が 1 のみ大きい場合 ($m+1$) と比較した値が ApEn である ($\text{ApEn}(n, m, r) = -\Phi^{m+1}(r) - (-\Phi^m(r)) = \Phi^m(r) - \Phi^{m+1}(r)$)。周期的で規則正しい時系列データならば, m あるいは $m+1$ 次元で作成された再構成ベクトルは時間遅れ座標系での分布が同じ性質なので, $\text{ApEn}(n, m, r) = \Phi^m(r) - \Phi^{m+1}(r) \approx 0$ となり規則正しい状態を判定できる。逆に乱雑な時系列データの場合, 値は大きくなる。生体信号から具体的に ApEn を計算したとする。ApEn は下限が存在し下限値は 0 で

ある。サインカーブやトーラスのような規則正しい時系列データならば 0 に近い値をとる¹⁶⁾¹⁷⁾。全く乱雑な場合が上限値となる。ApEn の上限値は時系列データの個数 n に依存する。乱雑な時系列データとするとデータ数が 600 個で ApEn は 1.4 前後, 2,000 個で ApEn は 1.9 前後となる¹⁶⁾¹⁷⁾。

また, 時間遅れの τ を導入して拡張した ApEn も報告されている¹⁹⁾。以下に式を示す。

$$\text{ApEn}(n, m, r, \tau) = \Phi^m(r) - \Phi^{m+1}(r)$$

$$= (n-(m-1)\tau)^{-1} \cdot \sum_{i=1}^{n-(m-1)\tau} \ln C_{m_i}(r) - (n-m\tau)^{-1} \cdot \sum_{i=1}^{n-m\tau} \ln C_{m+1_i}(r)$$

$C_{m_i}(r)$ は以下で定義される。

$$C_{m_i}(r) = (n-(m-1)\tau)^{-1} \cdot \sum_{j=1}^{n-(m-1)\tau} \theta(r-|X_i - X_j|)$$

$|X_i - X_j|$ はベクトル X_i とベクトル X_j の距離である。

2. Theiler's sigmas

Original data から計算された ApEn 値を Q_D とする。Original data から N 個の surrogate data を作成し, 各々 ApEn 値を求める。これらの ApEn 値が正規分布をとると仮定する。また, N 個の surrogate data から各々計算した ApEn 値の標本平均を μ_H , 標本標準偏差を σ_H とする。通常, iteratively refined surrogate を用いて時系列データの非線形性検定を行う。Theiler は次式で定義する検定統計量を “sigmas” とした⁸⁾。

$$\text{Sigmas} = \frac{|Q_D - \mu_H|}{\sigma_H}$$

ApEn 値が正規分布をとると仮定すると, sigmas > 1.96 であれば有意水準 $\alpha_D=0.05$ で, sigmas > 2.58 であれば有意水準 $\alpha_D=0.01$ で帰無仮説を棄却できる。Theiler は作成した surrogate data 数が少なくても sigmas を評価できるように t 分布を仮定した。即ち, 自由度 $(N-1)$ の t 分布を考える。作成した surrogate data が 10 個 ($N=10$) ならば t 分布表から $t_{0.05/9} = 2.26, t_{0.01/9} = 3.25$ なので, sigmas > 2.26 ならば有意水準 $\alpha_D=0.05$ で sigmas > 3.25 ならば有意水準 $\alpha_D=0.01$ で帰無仮説を棄却できる⁷⁾⁸⁾。

IV. 臨床応用例

図 5 は健常者の覚醒・睡眠中の呼吸運動をチェストバンド (RespiBand, NonInvasive Monitoring Systems Inc., USA) を用いて, inductance plethysmograph (Respi-somnograph, NonInvasive Monitoring Systems Inc) に信号を入力し A/D コンバーターで 10 Hz (0.1 秒) 間隔の

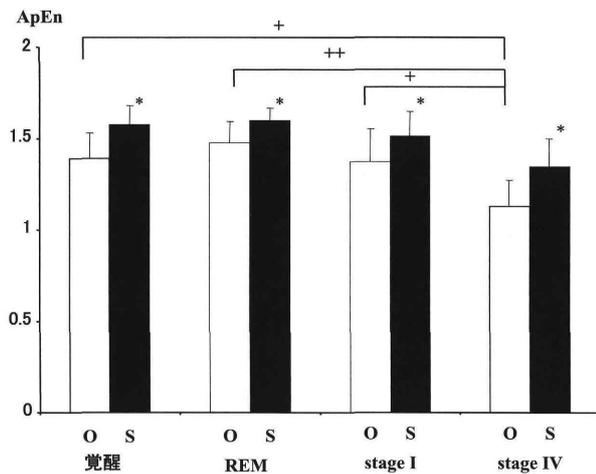


図5 覚醒時, 睡眠中の呼吸運動の時系列データ(original data)と surrogate data から計算した approximate entropy

閉眼覚醒, REM 睡眠, stage I 睡眠, stage IV 睡眠時の呼吸運動を 8 名の健常者から測定し, 3 分ごとの ApEn を計算した。Surrogate data (S) は各 10 個ずつ作成し平均値を利用した。Stage IV 睡眠時の呼吸運動は他よりも有意に ApEn が低値であった。各睡眠・覚醒段階の surrogate data の平均は original data (O) よりも有意に高値であった ($n=8$, mean \pm SD)。ANOVA ; $p<0.0001$; $+p<0.01$, $\#p<0.001$ (Scheffé's test) ; $*p<0.01$ (Wilcoxon matched-pairs signed-rank test)

(Burioka N, et al¹⁶). *Chest* 123 : 2003 より引用)

時系列データ(original data)にした後, ApEn を計算した結果である。全体として覚醒時, stage I, stage IV, REM 睡眠の ApEn 値は有意に変動していた¹⁶。特に深睡眠の stage IV で ApEn 値は有意に低値で, 他の睡眠覚醒段階より規則正しい呼吸運動になっていることが分った。また, 各 original data から 10 個の iteratively refined surrogate を作成し, Theiler's sigmas を計算した。作成した surrogate data が 10 個と少なかったため, t 分布表から $t_{0.05/2}(9)=2.26$ を利用して $\text{sigmas}>2.26$ ならば有意水準 $\alpha_0=0.05$ で帰無仮説を棄却できると考えた⁷⁾⁸⁾。

ApEn は乱雑な時系列データでは値が大きくなる⁵⁾¹⁸⁾。Surrogate data は作成過程で乱数を使用しているため, surrogate data の ApEn 値が original data の ApEn 値より大きな値の場合, original の時系列データは線形確率過程と異なると予想される(図5)。図6に示すように大多数の各睡眠・覚醒段階の呼吸運動で Theiler's sigmas > 2.26 になり「測定された生体信号の時系列データは線形確率過程から得られたものである」という帰無仮説が棄却された。したがって, 覚醒, 睡眠中に測定された呼吸運動

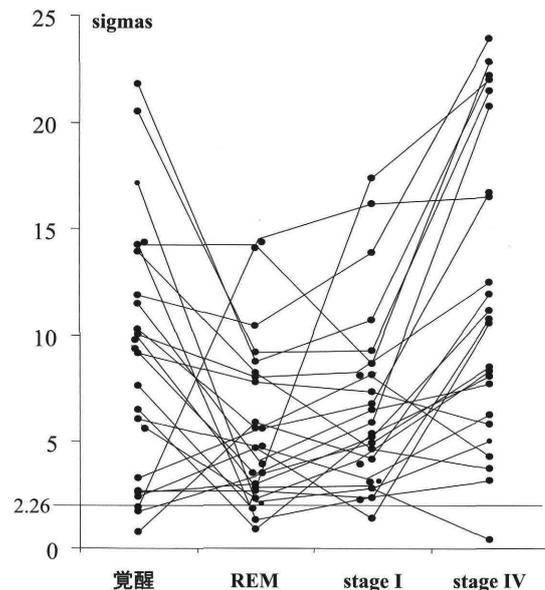


図6 覚醒時, 睡眠中の呼吸運動の非線形性検定

大多数の各睡眠・覚醒段階の呼吸運動で Theiler's sigmas は $\text{sigmas}>2.26$ になり「測定された生体信号の時系列データは線形確率過程から得られたものである」という帰無仮説が棄却された。黒丸は覚醒時, stage I, stage IV, REM 睡眠の呼吸運動の時系列データ化して Theiler's sigmas を計算した結果であり, 各被験者の値を線で結んだ。

(Burioka N, et al¹⁶). *Chest* 123 : 2003 より引用)

は非線形性を認める場合が多く, 特に深睡眠の stage IV で非線形性があると考えられた¹⁶⁾。

図7は① 健常者の安静閉眼覚醒時脳波, ② 欠神てんかん患者のてんかん発作のないときの脳波, ③ てんかん発作時の脳波を A/D コンバーターで 200 Hz (0.005 秒) 間隔の時系列データ(original data)にした後, ApEn を計算した結果である¹⁷⁾。欠神てんかん発作時の脳波から計算された ApEn は有意に小さく, より規則正しい波形になっていたと考えられた。各 original data から 40 個の iteratively refined surrogate data を作成し, ApEn を計算して Theiler's sigmas を求めた。図8に示すように健常者の安静閉眼覚醒時脳波および欠神てんかん発作がない時期の脳波の大多数は Theiler's sigmas ≤ 1.96 で帰無仮説が棄却されなかった。即ち, 健常者の安静閉眼覚醒時脳波および欠神てんかん発作がない時期の脳波は非線形性を認めないと考えられた¹⁷⁾。通常の脳波に非線形性を認めない理由として, 頭皮の表面から測定する脳波にはノイズが多い点と脳の複数領域からの電気信号が重なり合ってしまう点にあると推測された。したがって, 健常者の脳波は決定論的カオスから生じた生体信号であるという報告²⁰⁾には問題

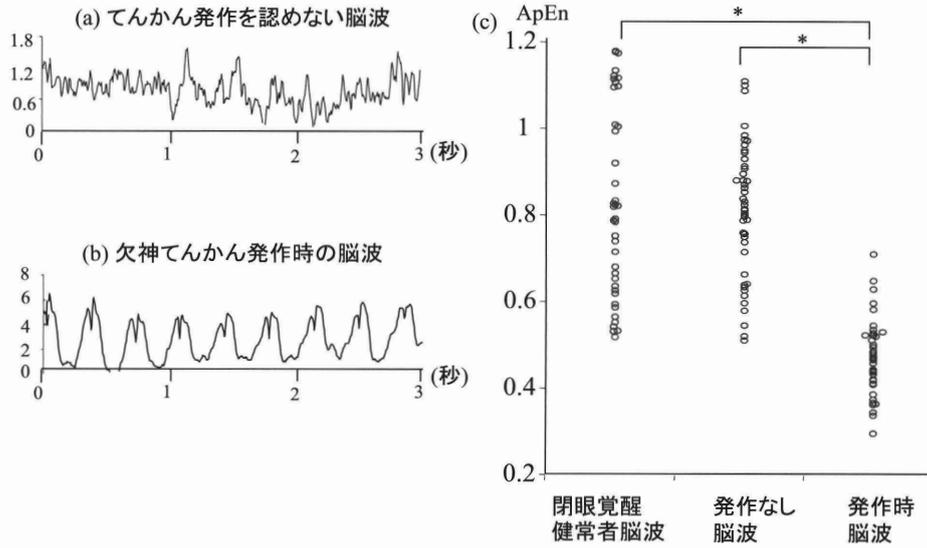


図7 健常者の安静閉眼覚醒時脳波および欠神てんかん患者の欠神てんかん発作のないとき、てんかん発作時の脳波から計算した approximate entropy
 a : 欠神てんかん患者のてんかん発作がない脳波部分。振幅が小さく、速波周波数のβ波である。
 b : 欠神てんかん患者のてんかん発作脳波部分。棘徐・徐波で波形に特徴があり振幅が大きい。
 c : 計算したすべての ApEn 値を使用した。安静閉眼覚醒時脳波 (n=40), 欠神てんかん患者の欠神てんかん発作のない脳波 (n=50), てんかん発作時の脳波 (n=50)。欠神てんかん発作時の ApEn は有意に低下していた (*p<0.01)。
 (Burioka N, et al¹⁷⁾. Clin EEG Neurosci 36 : 2005 より引用)

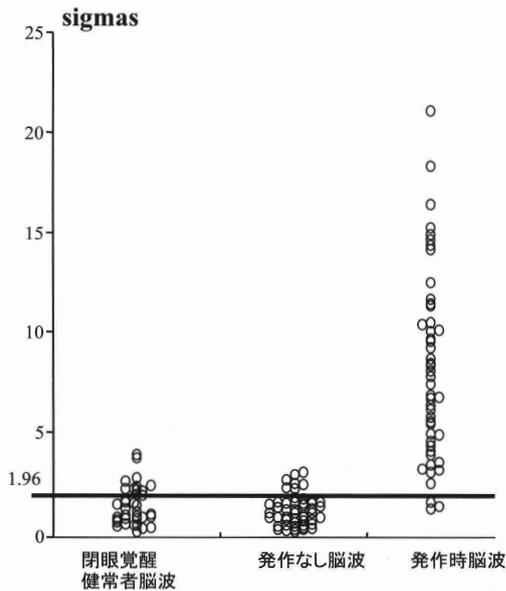


図8 健常者の安静閉眼覚醒時脳波, 欠神てんかん患者のてんかん発作のない脳波, てんかん発作時脳波の非線形性検定
 欠神てんかん発作時脳波の Theiler's sigmas は測定した発作時脳波のうち 94 % が sigmas > 1.96 になっていた。即ち、欠神てんかん発作時に脳波の非線形性が明らかになったと推測された。
 (Burioka N, et al¹⁷⁾. Clin EEG Neurosci 36 : 2005 より引用)

があると思われた。一方、欠神てんかん発作脳波の Theiler's sigmas は測定した発作時脳波のうち 94 % (47/50) が sigmas > 1.96 になっていた(図8)。また、50個の original data (欠神てんかん発作時の脳波の時系列データ)から各々40個ずつ作成した surrogate data の平均 ApEn は、

original data から算出した ApEn よりも大きな値となっていた¹⁷⁾(図9)。これらの結果から、欠神てんかんは脳の特定領域の神経細胞が同期性に異常放電して非線形システムとして規則性のある強い信号を発生し、測定された欠神てんかん脳波に非線形性を認めたと推測された²¹⁾。

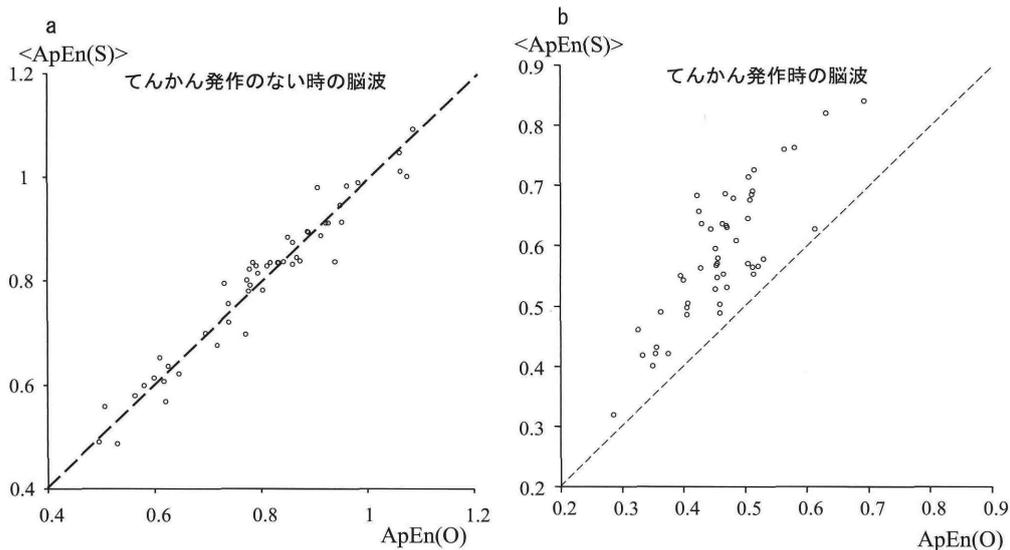


図9 欠神てんかん発作の有無による脳波 original data の ApEn および surrogate data から計算した ApEn の関係

各 original data から 40 個の iteratively refined surrogate data を作成し、各々 ApEn を計算した。図は original data の ApEn (O) と対応する各 40 個の surrogate data の ApEn 平均値、 $\langle \text{ApEn}(\text{S}) \rangle$ の散布図である。てんかん発作を認めない脳波 (a) では ApEn (O) と $\langle \text{ApEn}(\text{S}) \rangle$ は似た値であるが、てんかん発作脳波 (b) の各 original data から作成した対応する surrogate data の平均 $\langle \text{ApEn}(\text{S}) \rangle$ は、original data の ApEn (O) よりも大きいことが分る。

(Burioka N, et al¹⁷⁾. *Clin EEG Neurosci* 36:2005 より引用)

おわりに

Surrogate data の概念、種類と計算法、生体信号解析に応用した例を解説した。測定した時系列データの非線形性を統計学的に検定する概念は医学分野では新しいが、生理学的現象を新しい観点から展望できる可能性がある。一方、surrogate data 法で時系列データの非線形性を認めても決定論的カオスとは結論できない点に注意が必要である。決定論的カオスならばシステムから生じた信号は必ず非線形性を認めるが、逆は真とは限らないためである。

通常、Theiler's sigmas を用いて両側検定を行うので、surrogate data から計算した非線形ダイナミクスの特徴を表す統計量の平均値 (μ_H) が original data から計算された値 (Q_D) よりも大きいか小さいかも重要である。相関次元、ApEn、sample entropy は surrogate data から計算した値の平均値は original data から計算された値より大きくないと有意でも意味ある結果とならない。逆に最大リアプノフ指数はカオスであれば正の値をとり、規則正しければ 0 となり乱雑であればマイナスの値をとるため²²⁾、surrogate data から計算した最大リアプノフ指数の平均

値 (μ_H) は original data から計算された最大リアプノフ指数 (Q_D) よりも小さくしなければ Theiler's sigmas が有意水準を上回る値でも意味ある結果とならない⁷⁾。

生体信号から相関次元や最大リアプノフ指数を計算し生体信号はカオスであるとした報告の多くは、surrogate data 法が行われていなかったり、surrogate data から計算した相関次元の平均値が original data から計算された相関次元よりも小さくなっていたり²³⁾、surrogate data から計算した最大リアプノフ指数の平均値が original data から計算された最大リアプノフ指数よりも大きくなっており問題がある。今後、報告されたなかで surrogate data 法が行われている結果の再検討から生体信号と決定論的カオスの関係が明らかになるかもしれない。

Surrogate data 法は有用であるか限界と問題点が指摘されている²⁴⁾。そして、利用にあたっては注意が必要である。即ち、示したいのは「ダイナミクスに非線形性があるか」であるが、実際に行っているのは「時系列データが線形でない」ことを示すために「時系列データは線形確率過程である」という帰無仮説を棄却する点である。これはヒトから得られた生体信号を元システムがもつダイナミクスからの射影と仮定すると、ある程度の間接的な検証にはな

るか解釈を十分に考える必要がある。今後のさらなる開発と改良が期待される。

文 献

- 1) 北川源四郎. 時系列解析入門. 岩波書店. 東京: pp1—38, 2005
- 2) 日野幹雄. スペクトル解析. 朝倉書店. 東京: 2010
- 3) Grassberger P, Procaccia I. Characterization of strange attractors. *Phys Rev Lett* 50 : 346—349, 1983
- 4) Wolf A, Swift JB, Swinney HL, Vastano JA. Determining Lyapunov exponents from a time series. *Physica D* 16 : 285—317, 1985
- 5) Pincus SM. Approximate entropy as a measure of system complexity. *Proc Natl Acad Sci USA* 88 : 2279—2301, 1991
- 6) Richman JS, Moorman JR. Physiological time-series analysis using approximate entropy and sample entropy. *Am J Physiol Heart Circ Physiol* 278 : H2039—2049, 2000
- 7) 池口徹 山田泰司 小室元政. カオス時系列解析の基礎と応用. 第2刷. 合原一幸, 編. 産業図書株式会社. 東京: pp253—324, 2002
- 8) Theiler J, Eubank S, Longtin A, *et al.* Testing for nonlinearity in time series : the method of surrogate data. *Physica D* 58 : 77—94, 1992
- 9) Schreiber T. Interdisciplinary application of nonlinear time series methods. *Physics Reports* 308 : 1—64, 1999
- 10) Shreiber T, Schmitz A. Surrogate time series. *Physica D* 142 : 346—382, 2000
- 11) Takens F. Detecting strange attractors in turbulence. In : Rand D and Young BS, eds. *Lecture Notes in Mathematics* Vol. 898, Berlin. Springer : pp366—381, 1981
- 12) 川嶋弘尚 酒井英昭. 現代スペクトル解析. 森北出版. 東京: pp129—134, 2007
- 13) Schreiber T, Schmitz A. Improved surrogate data for nonlinearity tests. *Phys Rev Lett* 77 : 635—638, 1996
- 14) Ramdani S, Bouchara F, Julien Lagarde J. Influence of noise on the sample entropy algorithm. *Chaos* 19 : 013123, 2009
- 15) Nagarajan R. Surrogate testing of linear feedback processes with non-Gaussian innovations. *Physica A* 366 : 530—538, 2006
- 16) Burioka N, Cornélissen G, Halberg F, Kaplan DT, Suyama H, Sako T, Shimizu E. Approximate entropy of human respiratory movement during eye-closed wake and different sleep stages. *Chest* 123 : 80—86, 2003
- 17) Burioka N, Cornélissen G, Maegaki Y, Halberg F, Kaplan DT, Miyata M, Fukuoka Y, Endo M, Suyama H, Tomita Y, Shimizu E. Approximate entropy of the electroencephalogram in healthy awake subjects and absence epilepsy patients. *Clin EEG Neurosci* 36 : 188—193, 2005
- 18) Pincus SM. Approximate entropy as a measure of irregularity for psychiatric serial metrics. *Bipolar Disord* 8(5 Pt 1) : 430—440, 2006
- 19) Kaffashi F, Foglyano R, Wilson CG, Loparo KA. The effect of time delay on approximate & sample entropy calculations. *Physica D* 237 : 3069—3074, 2008
- 20) Pritchard WS, Duke DW. Measuring “chaos” in the brain : a tutorial review of EEG dimension estimation. *Brain Cogn* 27 : 353—397, 1995
- 21) Stam CJ. Nonlinear dynamical analysis of EEG and MEG : review of an emerging field. *Clin Neurophysiol* 116 : 2266—2301, 2005
- 22) 長島弘幸 馬場良和. カオス入門. 現象の解析と数理. 培風館. 東京: pp74—118, 1992
- 23) Kannathal N, Acharya UR, Lim CM, Sadasivan PK. Characterization of EEG—a comparative study. *Comput Methods Programs Biomed* 80 : 17—23, 2005
- 24) Kugiumitzi D. Evaluation of surrogate and bootstrap tests for nonlinearity in time series. *Studies in Nonlinear Dynamics & Econometrics* 12 (1) : ISSN (online) 1558—3708, 2008