平成 25 年度　未来研究ラボシステム　研究成果報告書

研究種目：展開研究　　研究期間：平成２５年１０月～平成２６年３月

研究課題名：情報論的階層性に注目したビッグデータ解析基盤の構築と応用

ラボ長

所属：機能創成・生体工学

氏名：清野 健

研究成果（当初の研究目的と得られた結果を記載ください。図表を含め２ページ程度）：

本研究の目的は，大規模データが有する情報論的構造の解析技術を開発し，生体信号データを具体例としてその有用性を示すことである．課題名にある情報論的階層性とは，情報理論もしくは統計力学における最大エントロピー原理から導かれるようなミクロな詳細に依存しないマクロ構造や法則性の出現を意味する．

本研究では，専門分野の異なる研究者間で問題意識を共有しつつ，個々の専門性を活かした研究体制を整えるために，本研究期間において２回の研究会を開催した．第１回研究会は平成 11 月 19 日に開催し，清野が心拍変動の大規模データ分析における問題点を解説した．第２回研究会は 2 月 20 日に開催し，鎌谷が混合モデルについて，加嶋がフィードバック制御系における確率雑音について，それぞれ解説した．これらの研究会での議論を参考に研究を進め，以下の成果をえた．

心拍変動の大規模データベースの分析

現在，国内外において大規模に医療情報を集積し，医療の質の向上，効率化に役立てようとする取り組みがはじまっている．大規模医療データの２次利用には，予測・予防医療に役立つ知識の発見，エビデンスに基づいた研究開発，医療政策決定など，発展的応用の可能性があり，その利活用のためのルール・法整備が検討されている．そのような医療大規模データの利活用の可能性と具体的な方法論を示すために，本研究ではホルター心電図から抽出された心拍変動時系列の大規模データを使った研究を実施している．従来の医療情報研究と異なる本研究の特徴は，時系列の大規模データを扱う点である．時系列データを読み解くためには，そこに内包された意味を抽出する解析手法が必要となる．

本研究では，約２０万例の心拍変動時系列に対して，従来用いられてきた時間領域指標，周波数領域指標を計算した．さらに，最近提案された心拍変動指標についても評価した．今回計算した指標は，心室性期外収縮時の変動パターンの評価する heart rate turbulence (HRT),

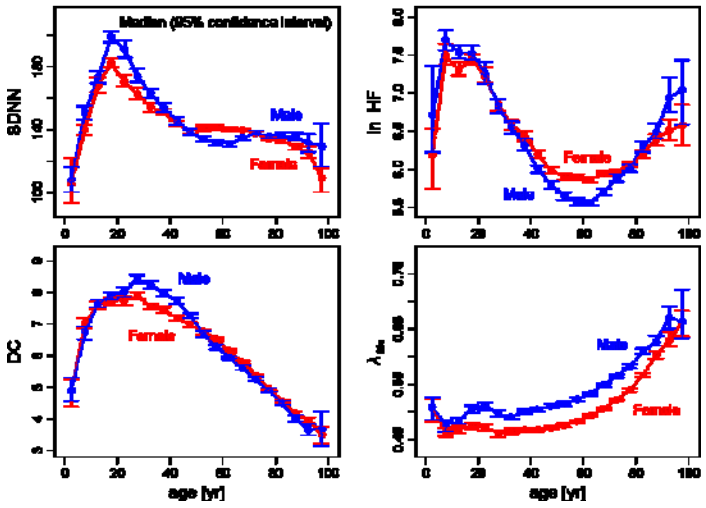

図 1: 心拍変動指標に見られる加齢と性差の影響．5歳階級ごとに中央値をプロットした．SDNN は心拍変動時系列(HRV)の標準偏差(ms)，HF は HRV の高周波領域(0.15–0.4 Hz)のパワー(ms^2)を表す．

心電図 RR 間隔変動の減少をアンカーとして変動パターンを評価する deceleration capacity (DC), 時系列の長時間相関を定量化する detrended fluctuation analysis を使ったスケーリング指標 α_1, 時系列の分散の非一様性を観測される非ガウス分布の構造に基づき定量化する非ガウス性指標 λ である．大規模データを扱うことで，加齢，性差と心拍変動指標の関係を詳細に調べることが可能になる（図 1）．とはいえ，このような分析は古典的なアプローチに過ぎない．今後，本研究では，気象，災害などの環境要因が心拍変動特性に及ぼす影響を，心拍変動の大規模データが有する時空間的構造に注目して分析する．

正規・ワイブル混合分布のパラメタ推定法の開発

心拍変動指標をはじめとして，生体信号時系列を特徴づける指標の確率分布のモデルとして，正規・ワイブル混合分布を仮定し，そのパラメタの推定法を開発した．正規・ワイブル混合分布モデルにおいて，正規分布は最大エントロピー原理に基づき推論されるものである．機能的異常のない生体システムおいて生体信号の特性は，それを評価する指標の代表値と散布度により特徴づけられること以外に事前に知ることのできる情報はないと仮定する．この仮定の下，最大エントロピー原理から正規分布が導かれる．一方で，ワイブル分布は故障現象のモデルとして広く使われているものである．生体システムにおいて特定の機能異常が生じ，それが不可逆な慢性的異常となった場合，その機能を特徴づける指標には適正値からの乖離が生じると考えられる．このような異常が現れる指標値の分布を，ワイブル分布でモデル化する．

正規分布とワイブル分布が共存する状態は，ヒトのような生体システム以外では考えづらい．なぜなら，生物ではない機械や装置では機能異常はシステムの停止に直結するため，停止したシステムと正常なシステムを混合して調べる状況はないからである．一方で，生体システムであれば，一部の機能異常（例えば，高血圧や脂質異常）が存在しても，システム全体としての生命は維持されるため，ある特定の機能については正常状態と異常状態が混在する可能性がある．さらに，正常値に自己修復しようとする作用も生体システム以外では存在しない．このような状況を記述するために，正規・ワイブル混合分布は我々が新たに導入したものである．

我々は，観測データに対し正規・ワイブル混合分布のあてはめるために，修正最尤法と EM アルゴリズムを用いる方法を開発した．ワイブル分布の位置パラメタについてはその尤度方程式が解を持たないため，EM アルゴリズムの通常の処方箋に基づき混合分布のパラメタを推定することはできない．そこで，本研究ではモデル全体の尤度が局所的に最大となるパラメタを修正最尤法で決定した．この方法の妥当性を，数値実験により検証するとともに，実際の心拍変動指標に対してもモデルのあてはめを行った．図 2 に示されるように，正

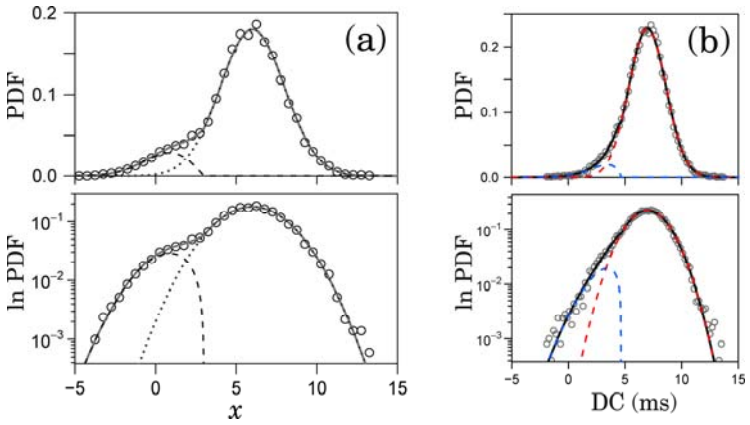

図 2: 正規･ワイブル混合分布のあてはめ．(a) 数値例．(b) 心拍変動指標 DC の分布．正規･ワイブル混合分布は，実際の指標分布に対し良い近似を与えた．

規・ワイブル混合分布は，実際の心拍変動指標の分布のモデルとして良い近似を与えることが確かめられた.

キーワード：
情報理論，最大エントロピー原理，生体情報，ベイズ推定，確率制御システム

研究経費（H25 年度）の内訳

備品費	消耗品費	旅費	謝金	その他	合計
228,140 円	248,075 円	258,160 円	65,625 円	0 円	800,000 円

共同研究者等

(1)共同研究者（氏名・所属）
加嶋 健司・システム創成専攻システム科学領域（現在，京都大学大学院情報学研究科）
鎌谷 研吾・システム創成専攻数理科学領域
(2)研究協力者（氏名・所属・学年（学生の場合））
梅津 佑介・システム創成専攻システム科学領域・M１
河田 洋輔・システム創成専攻システム科学領域・M１
東 晃弘・機能創成専攻生体子工学領域・M１
遠藤 哲太朗・機能創成専攻生体子工学領域・M１
竹内 公一・機能創成専攻生体子工学領域・M１

発表論文等（平成 26 年 3 月 31 日現在）

〔雑誌論文〕

[1] Kengo Kamatani, “The order of degeneracy of Markov chain Monte Carlo method”, Journal of the Japan Statistical Society, accepted, 2014
[2] 東　晃弘,清野　健,野村　泰伸,「正規・ワイブル混合分布モデルを用いた潜在情報の抽出」, 統計数理研究所共同研究リポート，出版予定.

〔著書〕

〔学会発表〕

[1] Kengo Kamatani, Local consistency of Markov chain Monte Carlo with some applications, MCMSki IV, Jan 6-8, Chamonix, France, 2014.
[2] Ken Kiyono, Non-Gaussian characteristics of heart rate variability in health and disease, Multifractal Analysis: From Theory to Applications and Back, Feb 27, Banff, Canada, 2014.
[3] 遠藤哲太朗，藤井翔太，清野健,「Mellin 変換を用いた間欠性ゆらぎの特徴付け：心拍変動解析への応用」，日本物理学会第 69 回年次大会，東海大学，神奈川県，２０１４年３月２０日.

〔その他〕

外部資金獲得状況・申請状況（本研究課題に関連して、科研費、JST 等の競争的資金、受託研究、奨学寄付金を受給された場合、また、申請された場合はその状況を記入ください）

科研費基盤研究（B）（一般）（H26～H28）へ申請した.

参考となるＨＰ等

http://www.sigmath.es.osaka-u.ac.jp/~kamatani/research/