

研究種目：基盤研究（C）  
研究期間：2007～2010  
課題番号：19500247  
研究課題名（和文） 混獲データの解析

研究課題名（英文） Analysis of Bycatch Data

研究代表者

南 美穂子 (MINAMI MIHOKO)  
慶應義塾大学・理工学部・教授  
研究者番号：70277268

研究代表者の専門分野：統計科学

科研費の分科・細目：情報学・統計科学

キーワード：国際研究者交流，分布の回帰分類木，海洋資源，一般化主成分分析，Impurity Measure, Kullback-Leibler 情報量，ゼロの多いデータ，Tweedie 分布

### 1. 研究計画の概要

混獲データは、対象魚種の漁獲データとは異なる特徴を持つ。これは操業の場所や条件はあくまでも対象魚種を獲るために選択されるのであって（一部の海洋哺乳類を除いては）混獲生物の状態とは直接の関係がないことによる。混獲データの数値上の特徴はゼロの割合が高いことである。本研究では全米熱帯マグロ類委員会 (IATTC) の Cleridy Lennert・Cody 博士と共同で混獲生物の資源評価・生態学的考察のための統計手法の開発とデータ解析を行う。また、海洋資源評価のための一般的な統計手法の開発も行う。

(1) **混獲生物の資源動向の解析のための統計手法の研究**：サメやウミガメといった個別の混獲生物の資源動向を把握するための統計手法を開発する。漁獲対象魚種に対する CPU 解析に対応するものであり、様々な条件（説明変数）が与えられたときの混獲数の条件付平均を推定するという回帰問題である。これまでに、平滑法を用いた Zero-inflated 負の 2 項回帰モデルを提案し、浮遊物によるマグロ巻網漁のサメの混獲数を解析した研究を行っているが、平滑パラメータの選択方法、係数の推定方法、モデル選択の方法についてより良い方法を考え、提案する。

(2) **混獲組成の解析**：各漁で混獲された様々な海洋生物の混獲数データから海洋生物種間のかかわりを解析するために、非正規性の強いデータから特徴量を抽出する方法を開発する。これは視点をかえると非正規データの次元の削減問題と捉えることができる。非負行列因子分解法 (Non-Negative Matrix

Factorization, NMF) の統計学的見地からの改良、具体的には Tweedie 分布、一般化線型モデルの概念を NMF に用いた方法の研究を行う。

(3) **対象魚種データの解析手法を混獲データに応用することによる問題点の紹介**：対象魚種データの解析に用いられている手法を性質の異なる混獲データに応用する場合に起こりうる問題点を統計理論の視点から指摘する。

### 2. 研究の進捗状況

(1) **混獲生物の資源動向の解析のための統計手法の研究**：平滑法を用いた Zero-inflated 負の 2 項回帰モデルによる、浮遊物によるマグロ巻網漁のサメ混獲数の解析については課題を解決し、論文にまとめたものが、学術雑誌 Fisheries Research に掲載された。

(2) **混獲組成の解析**：海洋生物種間のかかわりを解析するための統計手法として一般化主成分分析法 (GPCA 法) を提案した。この手法は、一般化線型モデル (GLM) が線型モデルを拡張することと同様の意味において主成分分析法を拡張するものである。また特に、ゼロの割合が高く非負の値を取るという漁獲・混獲データの特徴に特に適したものとして Tweedie 分布を用いた Tweedie-GPCA 法を提案した。この手法を用いて東部太平洋でのマグロ巻網漁による漁獲・混獲データを解析したところ 4 つの特徴量で 70% の尤離度 (deviance) を説明でき、また、陸からの距離や海流などの環境要因との関係が強い特徴量が得られた。結果は国際学会、国際ワー

クシヨップで発表した。この手法は混獲データに適したものであるが、遺伝学や他の分野での適用も考えられる。一般的な設定でのシミュレーションを計画し今後論文にまとめる予定である。

(3) **対象魚種データの解析手法を混獲データに応用することによる問題点の紹介**：負の2項回帰モデルをゼロの割合の多いデータに用いると減少（あるいは増加）傾向を極端に示し資源評価には問題が大きいことを実データとシミュレーションデータの両方で示し、理由を解析した。結果を国際学会などで発表し、現在、論文作成中である。

(4) **体長分布の空間的要因・季節・環境要因による回帰分類法**：海洋生物の体長は場所・季節などによってその分布が異なる。体長の分布を推定・分類することは海洋生物の特性について推測をする上で、また資源管理の観点からも重要である。本研究はこれまで便宜的な方法が用いられてきたこの問題に統計科学的な手法を提案しようとするものである。Kullback-Leibler 情報量を不純度に用いた分布の回帰分類法を提案し、マグロ体長の頻度データに適用したものを論文として発表した。また、国際学会でも講演する予定である。この課題は当初の計画には含まれていないが、実際の資源評価に貢献度の高い重要なものであると考える。

### 3. 現在までの達成度

- (1) **混獲生物の資源動向の解析のための統計手法の研究**：課題はまだあるが重要な問題は解決し、論文が学術雑誌に掲載されたので②おおむね順調に進展している、と評価する。
- (2) **混獲組成の解析**：あらたな手法を考案し、国際会議などでの発表でも良い評価を得ており、②おおむね順調に進展している、と評価できる。
- (3) **対象魚種データの解析手法を混獲データに応用することによる問題点の紹介**：種々の減少を理論的に解明することができたので、②おおむね順調に進展している、と評価する。
- (4) **体長分布の空間的要因・季節・環境要因による回帰分類法**：当初の計画に含まれていないが、重要な課題であり、①当初の計画以上に進展している、と評価する。

### 4. 今後の研究の推進方策

- (1) **混獲生物の資源動向の解析のための統計手法の研究**：おおむね課題は達成したが、他の誘引法による巻き網漁のデータが入手可能であれば解析を試みたい。
- (2) **混獲組成の解析**：シミュレーションと論文作成を行う。
- (3) **対象魚種データの解析手法を混獲データに応用することによる問題点の紹介**：追加のシミュレーションを行い、論文を執筆する。

(4) **体長分布の空間的要因・季節・環境要因による回帰分類法**：体長データの入手を待って、密度推定を用いた方法で解析し、論文を執筆する。国際会議で発表し海外研究者と議論・情報交換する。

### 5. 代表的な研究成果

（研究代表者、研究分担者及び連携研究者には下線）

〔雑誌論文〕（計3件）

- (1) C. E. Lennert-Cody, M. Minami, P. K. Tomlinson, M. N. Maunder, Exploratory analysis of spatial-temporal patterns in length-frequency data: An example of distributional regression trees, Fisheries Research, 査読有, 102 巻, 2010, 323-326.
- (2) M. Minami, C. E. Lennert-Cody, W. Gao and Román-Verdesoto Modeling shark bycatch: the zero-inflated negative binomial regression model with smoothing, Fisheries Research, 査読有, 84 巻, 2008, 210-221

〔学会発表〕（計9件）

- (1) 南 美穂子, 分布に対するCART (分類・回帰木) : 体長分布の時空間によるパターン解析, 統計関連学会連合大会, 2009年9月9日, 同志社大学 (京都市)
- (2) Mihoko Minami and Cleridy Lennert-Cody, A New Feature Extraction Method for Very Non-Normal? Data: Analysis of Multivariate Species-size Data from a Tuna Purse-seine Fishery, Workshop on Spatial analysis for stock assessment, 2008年10月14日, ラ・ホヤ (アメリカ合衆国)
- (3) Mihoko Minami, A New Feature Extraction Method for Very Non-Normal Data: Analysis of Multivariate Species-size Data from Purse-seine Tuna Fisheries, International Biometric Conference 2008, 2008年7月14日, ダブリン (アイルランド)
- (4) Mihoko Minami, Modeling Shark Bycatch: The Zero-Inflated Negative Binomial Regression Model with Smoothing, Joint Statistical Meeting 2007, 2007年7月28日, ソルトレークシティ (アメリカ合衆国)