

**Investigation of Goodness-of-fit
in Ecological Data Modeling**

July 2016

MAYUMI NAKA

主 論 文 要 旨

| | | | | |
|---|-----|---|----|------|
| 報告番号 | 甲乙第 | 号 | 氏名 | 仲 真弓 |
| 主論文題目： Investigation of Goodness-of-fit in Ecological Data Modeling (生態系データモデリングにおける分布の適合性の研究) | | | | |
| (内容の要旨) 本論文は、観測されたデータに対して、どのようにしたら適切なモデルを導くことができるか、その適合性はどのようにして検証したらよいのか、という観点から行ってきたこれまでの研究成果で構成されている。前半では、トロール漁の影響調査データと淡水流入による影響調査データに対して、外的要因の変化がどのように海底生物や河口でのバナナエビの生態に影響するかを説明する確率分布を用いたモデルを構築し、解析した結果を示している。後半では、前半でも用いられた分布の適合性の検証によく用いられる Cramér-von Mises 統計量の理論的な性質、特に最小距離推定量を用いてパラメータ推定したときのロバスト（頑健）性を明らかにする。 第一章では、確率分布を用いたモデルの適合性検証に用いられる方法について、既存の研究をまとめる。 第二章は、オーストラリア北部海域における海底生物へのトロール漁の影響調査データに対し、確率分布を用いたモデルを構築して解析した結果を示す。トロール漁前の海底生物の体重分布のモデルとして、成長のモデルとして知られる確率微分方程式の定常分布として得られるガンマ分布を考え、分布の適合度検定統計量の一つである Cramér-von Mises 統計量をサンプルが独立だが必ずしも同じ分布に従っていない場合に拡張した統計量を用いて分布の適合性を検証する。80 ケース中 57 ケースで、体重のデータのモデルとしてガンマ分布を用いて問題がないことを確認した上で、このモデルを用いてトロール漁後の体重分布が変化したかを検定することで、トロール漁による影響を種ごとに検証する。 第三章では、オーストラリアの河口における淡水流入による生物への影響を解析している。ここではバナナエビの捕獲データに着目し、観測時期の水温と塩分濃度の影響を加味した成長と生存のモデルの非対称混合分布を用いて、河口に滞在するバナナエビの甲殻の長さの分布のモデルを導出する。モデルの適合性については、離散分布に対する Cramér-von Mises 統計量によって、19 ケース中 15 ケースでは、導出した混合分布を甲殻の長さのデータのモデルとして用いて問題ないことが確認でき、この 15 ケースについては、淡水流入による水温・塩分濃度の変化がバナナエビの成長に与える影響がこのモデルで説明できていると考えられる。 第四章は、 Cramér-von Mises 統計量を用いた分布の適合性に対する検証について、少量の異質なデータが混ざっていた場合についての漸近的な性質を示している。まず、パラメータを既知としたときの Cramér-von Mises 統計量の漸近分布が、非心カイ二乗確率変数の重み付き和の分布で表せ、異質なデータの影響は非心度にのみあらわれることを示す。次に、パラメータが未知で、パラメータ推定に最小 Cramér-von Mises 距離推定量を用いた場合についても同様に漸近分布を導出する。さらに、シミュレーションによる実験から、パラメータ推定に最小距離推定量を用いることで Cramér-von Mises 統計量のロバスト性が確保できることを明らかにする。 | | | | |

SUMMARY OF Ph.D. DISSERTATION

| | | |
|--|-------------------------------|-------------------------------------|
| School Fundamental Science and Technology | Student Identification Number | SURNAME, First name NAKA, Mayumi |
| Title Investigation of Goodness-of-fit in Ecological Data Modeling | | |
| Abstract This thesis discusses how we build models for ecological data and how we examine their goodness-of-fit to the observations. This thesis consists of two parts. In the first part, we present two case studies where statistical models are built by considering theoretical aspects of the data as well as their statistical characteristics. The case studies in the first part are for the assessments of the effect of trawling on animals on seabed and of the effect of freshwater flows on banana prawns in an estuary. In the second part, we show several theoretical properties of the Cramér-von Mises goodness-of-fit statistic, which is used in the first part, especially the robustness of the statistic when the minimum distance estimator is used. In Chapter 1, a brief summary of the methods for examining goodness-of-fit of distributions is given. In Chapter 2, experimental survey data for the assessment of trawling effect in the Northern Prawn Fishery in Australia are investigated. Gamma distribution is adopted as weight distribution of animals on seabed before trawling because it can be derived as the equilibrium distribution of the stochastic growth model. The goodness-of-fit of the distribution is examined by an extended version of the Cramér-von Mises statistic for independent but not identically distributed observations. It is shown that the gamma distribution can be used for modeling the weights of animals before trawling for 57 cases out of 80. The effect of trawling is verified for each species with this model. In Chapter 3, survey data for the assessment of the effect of freshwater flows in the estuary in Australia are investigated. The model we employed for the binned carapace length of banana prawns in the estuary is an asymmetric mixture distribution derived by combining a growth model with temperature and salinity of water and a survival rate model. It is shown that the model is appropriate for 15 cases out of 19 by using the Cramér-von Mises statistic for discrete distributions. For these 15 cases, the model can explain the effects of the changes in temperature and salinity of water caused by freshwater flows to the growth of banana prawn. In Chapter 4, asymptotic behavior of the goodness-of-fit test statistic when contamination exists is investigated. The asymptotic distribution of the Cramér-von Mises statistic for contaminated data when parameters are known is derived as a distribution of a weighted infinite sum of non-central chi-squared random variables. The effect of the contamination appears only in the non-centralities. The asymptotic distribution when the parameters are unknown and are estimated by the minimum Cramér-von Mises distance method is also derived. The theoretical results together with the result of numerical experiments show that the Cramér-von Mises statistic becomes robust when we use the minimum distance estimator for the estimation of parameters. | | |