

# ベイジアンネットワークモデルを用いた 長崎県の地域別農業動向分析

池森恵子<sup>1)</sup>

キーワード：農業センサス，個票データ，農業動向分析，ベイジアンネットワーク，離農要因

A Case Study of Regional Agricultural Trend Analysis in Nagasaki Prefecture  
Using a Bayesian Network Model

Keiko IKEMORI<sup>1)</sup>

## 目次

1. 緒言
2. 分析の概要
  - 1) 分析対象とした地域および農家属性
  - 2) 分析データの概要
  - 3) 単純集計
  - 4) ベイジアンネットワーク分析
3. 地域の概要と単純集計結果
  - 1) 島原地域
  - 2) 県北地域
4. ベイジアンネットワークによる農業構造の可視化と確率推論
  - 1) 2地域のベイジアンネットワークモデル
  - 2) 確率推論による影響度評価
  - 3) 単純集計と確率推論結果に基づく解釈
5. まとめ
6. 摘要
7. 引用文献

Summary

---

<sup>1)</sup> 現長崎県農林部農業イノベーション推進室

## 1. 緒言

長崎県は九州の西北部に位置し、三方を海に囲まれ、年間平均気温は17～18℃、年間降水量は2,000mm前後で、対馬暖流の影響を受けて冬は比較的暖かく、夏は涼しい海洋性の気候に恵まれている。半島や島が多く、山や丘陵が大部分で平地に乏しい地形により、農林水産省（以降、農林水産省、2022a）によれば令和4年の耕地面積は45,700haで九州では沖縄に次いで小さく、耕地率は11%程度である。本県の農業生産額は全国20位の1,551億円（農林水産省、2022b）で、肉用牛やばれいしょ、果樹や施設園芸など地域の特色豊かな農業が行われている。

他方、令和2年国勢調査によれば本県の人口は1,313千人で、2015年～2020年の5年間で約65千人減少しており、自然減に加えて若年層の人口流出による社会減により高齢化率は33.0%と、全国平均28.7%と比較すると高い水準となっている。この人口動態の変化は農業にも多大な影響を与えている。2020年農林業センサス（以降、20年センサス）では、長崎県の農業経営体数は17,936経営体で、2015年センサスに比べ3,972経営体、率にして18.1%減少した。

近年、少子高齢化による人口構造の変化や、国や地方の財政赤字の拡大によって政策運営が難しくなっていることを背景に、客観的証拠に基づく政策立案（Evidence Based Policy Making, EBPM）への取り組みが推進されている。EBPMの政策サイクルにおいては最初に政策の必要性の根拠となるエビデンス、つまり統計データ等の分析による現状把握が重要であるとされる（大竹ら、2022）。これまでに、農業センサス個票データを用い、統計的手法によって地域別の離農要因分析や離農予測を行った事例は多数報告されている（細山、2012；安武、2018；寺谷、2019）。例えば栗原ら（2014）は、2択の選択肢に対する選択確率をモデル化するプロビットモデルによって営農継続農家と離農農家との属性の違いを明らかにしている。

本研究では、農業センサス個票データを用いて、複雑な事象間にある因果を可視化して表現するベイジアンネットワークモデルを構築し、営農継続農家と離農農家との属性の違いについて分析を行ったので、本報告に記す。

## 2. 分析の概要

### 1) 分析対象とした地域および農家属性

分析対象は本県南東部に位置する島原地域および県北部に位置する県北地域の2地域（図1）の家族



図1 分析対象地域

経営体とした。また、センサス農林業調査票（農林水産省、2015）の設問において農産物販売額1位の部門に耕種品目をあげた経営体を、ここでは耕種農家として分類し、分析対象を耕種農家に絞った。

### 2) 分析データの概要

本研究では15年センサスおよび20年センサス個票データを使用した。先行研究と同様に、15年センサスと20年センサスの個票データを接続し、ここで接続できた（15年センサスと20年センサスの両方に存在する）農家を‘営農継続’農家とした。接続できなかった（15年センサスに存在し、20年センサスからは消えている）農家について地域データと照合し、営農継続農家をさらに検索し精査を行った。精査においても接続できなかった農家を‘離農’農家として分類した。精査後の分析データ数および‘営

農継続’‘離農’農家の分類結果を表1に示す。

なお、本研究で使用したセンサス個票データは「農林業センサスに係る調査票情報の提供について（通知）」（農林水産省3統計第504号，令和3年6月24日）により利用許可を得ている。

表1 分析データ数および‘営農継続’‘離農’農家数（経営体数）

項目	分析データ数	継続農家数 (割合)	離農農家数 (割合)
島原地域	5,141	4,153 (80.8%)	988 (19.2%)
県北地域	4,202	3,223 (76.7%)	979 (23.3%)

### 3) 単純集計

ベイジアンネットワークモデルによる推論結果を理解するにあたり、2地域の15年センサスおよび20年センサスの個票データを単純集計し、基幹的農業従事者の平均年齢別、販売金額1位の品目別、農産物販売額別、および経営耕地面積規模別の割合を算出した。基幹的農業従事者の平均年齢は、センサス調査票において「ふだん仕事として主に農業に従事している」と回答した世帯員の年齢平均を世帯ごとに算出し、基幹的農業従事者が不在の世帯については‘基幹的農業従事者がいない’として区分した。また、販売金額1位の品目は同調査票の設問において、過去1年間に農産物の販売があり、かつ、農産物の販売額が1位と答えた耕種部門を集計したものである。なお、「農産物の販売がない」と答えた世帯については‘販売なし’として区分した。

### 4) ベイジアンネットワーク分析

ベイジアンネットワークは、複数の変数間の因果関係を表すグラフ構造と、個々の変数に割り当てられた条件付き確率分布群によって定義される確率モデルである（本村・佐藤，2000；本村，2003，繁樹ら，2006）。グラフ構造では、確率変数をノードで表し、そのノードの間をリンクという循環しない矢線でつなぎ、事象間の因果関係を表現する。例えば変数  $X_i$  と  $X_j$  があり、 $X_j$  が  $X_i$  に依存している関係を表すとき、 $X_i \rightarrow X_j$  と表す。このときの  $X_i$  を親ノード、 $X_j$  を子ノードと呼ぶ。

本研究におけるベイジアンネットワークモデルの構築には、ベイジアンネットワーク構築支援システム BayoLinkS（株式会社 NTT データ数理システム）を用いた。本ソフトには感度分析機能が搭載されており、構築したモデルにおける相互情報量の算出や確率推論に基づく事前確率値と事後確率値の差分によって説明変数と目的変数の依存程度を計量化できる。

モデル構築にあたっては、センサス個票データの調査項目のうち13項目を説明変数の候補とし（表2）、その一部には集計や平均値算出等の前処理を施した。また、先述した方法により‘営農継続’‘離農’の分類結果を目的変数（＝‘継続/離農’）に設定した。島原地域のモデル構築において採用した説明変数のデータカテゴリを表3に示す。変数値が量的変数である場合は、BayoLinkSのグルーピング機能を活用して最大6のカテゴリ変数へ変換した。さらに、説明変数は「経営形態」「経営基盤」「販売体系」の3グループに分割し構造学習を行った。なお、表1に示した地域ごとの全データのうち7割をベイジアンネットワークモデル構築の訓練用データとし、残りをモデル検証用データとした。

表2 バイジアンネットワーク分析の説明変数の候補

候補とした説明変数	島原地域	県北地域
基幹的農業従事者の平均年齢	○	○
基幹的農業従事者数	○	○
他産業従事者数（自営含む）	○	○
農業専従者数	×	×
所有する農地面積	×	×
貸地面積	×	×
借地面積	○	×
経営耕地面積	○	○
農産物販売額	○	○
販売金額第1位の営農類型	○	○
常時雇用の有無	×	×
臨時雇用の有無	○	×
第1位の出荷先	○	○

○：最終的に説明変数に採用した

×：最終的に説明変数に採用しなかった

表3 島原地域の分析カテゴリ

グループ	説明変数		カテゴリ				
経営形態	基幹的農業従事者の平均年齢	基幹的従事者がいない	50代以下	60代	70代	80代以上	—
	世帯あたり基幹的農業従事者数	基幹的従事者がいない	1人	2人	3人以上	—	—
	他産業従事者数（自営含む）	他産業従事者がいない	1人	2人以上	—	—	—
経営基盤	販売金額第1位の営農類型	販売なし	水稻	麦・雑穀・豆類	露地園芸	施設園芸	果樹類
	経営耕地面積	50a未満	50a以上100a未満	100a以上150a未満	150a以上	—	—
	借地面積	借地していない	50a未満	50a以上	—	—	—
	農産物販売額	販売なし	50万円未満	50万以上300万円未満	300万円以上1000万円未満	1000万円以上	—
販売体系	臨時雇用の有無	いない	いる	—	—	—	—
	第1位の出荷先	販売なし	農協	農協以外の市場出荷	直接取引	—	—

### 3.地域の概要と単純集計結果

#### 1) 島原地域

雲仙普賢岳を中心とする穏やかな丘陵地帯と海岸沿いに広がる平野部からなる島原地域は、周囲を海に囲まれた半島で、本県の農業算出額の約4割を占める県内随一の農業地帯である。雲仙火山に由来する肥沃な黒ボク土壌に恵まれ、長崎県全体の畑耕地面積の約3割を占める。

この地域の単純集計の結果を図2に示す。2015年～2020年において販売金額1位の品目別割合に大きな変化はなく、露地園芸、施設園芸および果樹類を合わせた園芸品目の割合が6割強を占めた。農産物販売額別割合は‘1000万円以上’で4.5ポイント上昇し、販売金額300万円以上の農家は53.1%と半数を超えた。また、経営耕地面積規模別割合は‘150a’以上で2.9ポイント上昇し、規模拡大による経営基盤の強化を図った経営体の動きがみられた。また、基幹的農業従事者の平均年齢割合は‘50代以下’および‘60代’の合計が依然として5割を占め、地域の半数の経営体で堅調な事業承継が行われていた。その一方で、‘基幹的農業従事者がいない’は7.0ポイント上昇し、農業が基幹産業である当該地域においても基幹的農業従事者不在経営体の顕著な増加が確認できた。

#### 2) 県北地域

県北地域は山や丘陵の起伏が海岸線まで迫る平坦地の少ない地形から、耕地が傾斜地に多く、棚田水田が発達した地域である。また、粘質の過乾過湿な土壌によって畑作物の作付けに不向きな耕地が多いなど不利な生産条件を多く抱える。

単純集計結果を比較したグラフを図3に示す。営農類型では、水稻主体が過半と、園芸品目主体が3割を占める構成にほとんど変化はみられなかった。経営耕地面積規模別では、‘150a以上’で1.9ポイントのわずかな上昇がみられたが、全体の構成にほとんど変化はみられなかった。農産物販売額別では‘販売なし’および‘50万円未満’を合計した割合は57.0%と、零細規模の農家割合が過半を占める一方で、販売額50万円以上の各カテゴリの割合はわずかに上昇し、‘50万～300万円未満’で最も高い6ポイントの上昇がみられた。基幹的農業従事者の平均年齢別では、この5年間で‘80代以上’が10ポイント近く下落した。一方で、本県の基幹的農業従事者の平均年齢である‘60代’で4.6ポイント下落し、‘70代’が依然として最も高い割合を占めていた。事業承継が進まないなか、高齢従事者のリタイアが進んでいる現状を捉えていた。さらに、基幹的農業従事者不在経営体の割合は16.6ポイント上昇しており、農業労働力の劣弱化が懸念された。

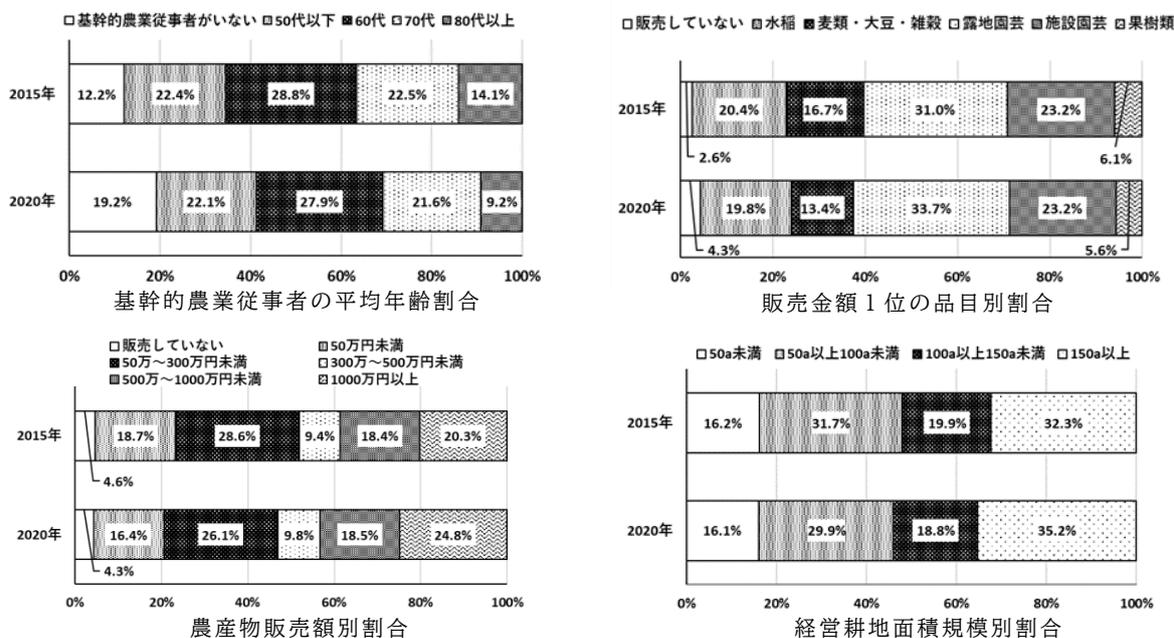


図2 島原地域におけるセンサ単純集計結果

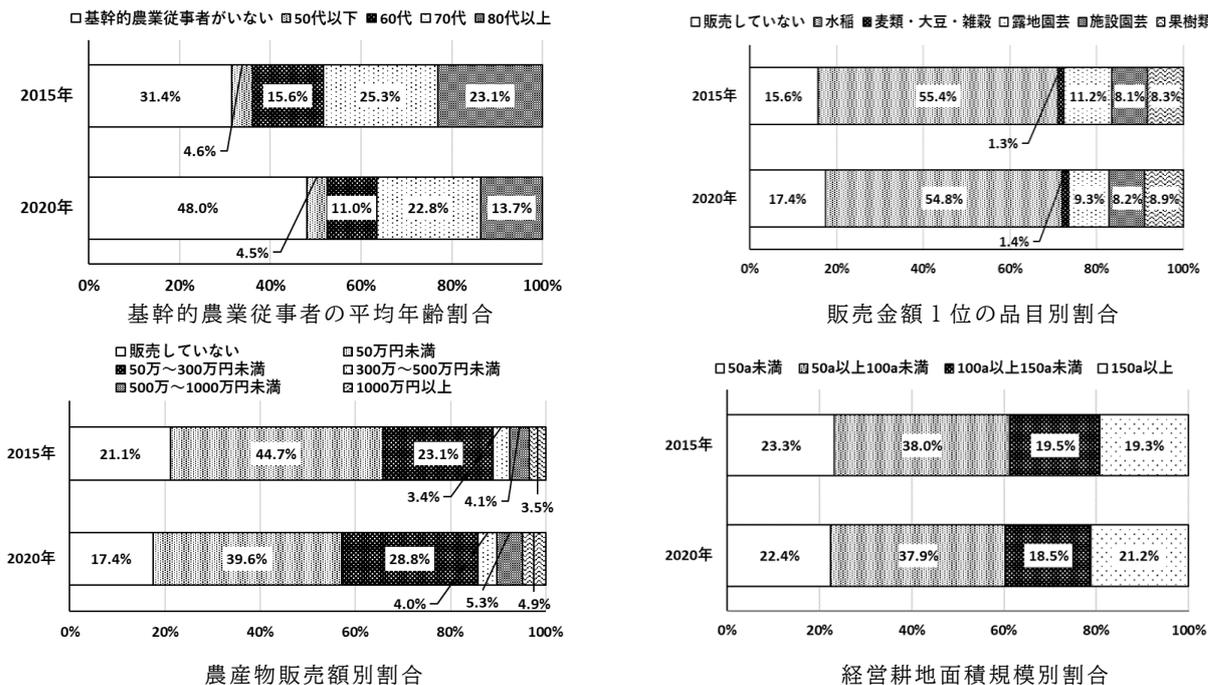


図3 県北地域におけるセンサス単純集計結果

#### 4. バイジアンネットワークによる農業構造の可視化と確率推論

##### 1) 2地域のバイジアンネットワークモデル

島原地域および県北地域のバイジアンネットワークモデルを図4および図5に示す。モデル検証用データを用いた評価の結果、島原地域および県北地域の正解率はそれぞれ83.7%と79.7%であった。

両地域のモデルを比較すると、採用されたノードに違いがみられた。県北地域では、「借地面積」と「臨時雇用の有無」の2つのノードが目的変数「継続/離農」との因果関係が認められず採用されなかった。島原地域のモデルの感度分析を行ったところ、借地や臨時雇用のある経営体は経営規模や販売金額の大きい経営体、あるいは、営農類型が園芸品目である経営体でみられた(データ略)。零細規模の経営体が過半を占め、水稻主体の県北地域において、これら2つのノードの採用されなかった理由は論理的に説明できる。

他方、両モデルに共通していたのは、基幹的農業従事者の平均年齢、「経営耕地面積」、「農産物販売額」の3つのノードが目的変数「継続/離農」に直接影響を与える親ノードとして示されたことであ

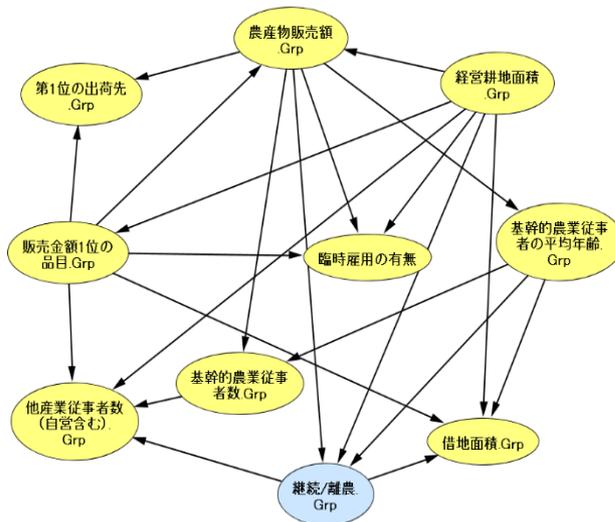


図4 島原地域のバイジアンネットワークモデル。まさに、経営資源「ヒト・モノ・カネ」の3要素の重要性を裏付けていた。

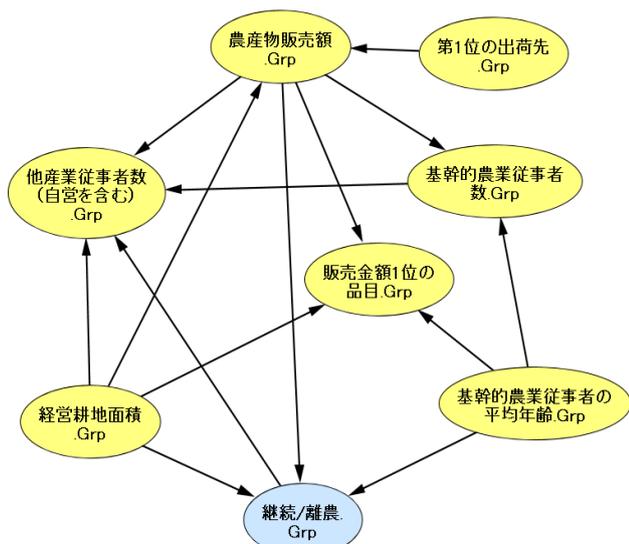


図5 県北地域のベイジアンネットワークモデル

## 2) 確率推論による影響度評価

これら3つの親ノードと目的変数「継続/離農」との関係性を評価するため、ノードごとに確率推論を実行する。確率推論とは、構築したベイジアンネットワークの条件付き確率分布に対して、一部の変数に観測値（証拠）を与えることでネットワーク全体の確率分布を更新し、ユーザーの指定する特定の条件下における各ノードの発生確率（事後確率）を求める計算処理である。目的変数「継続/離農」のノードに観測値として「営農継続」、 「離農」をそれぞれ与えることにより、営農継続しやすい、あるいは、離農しやすい傾向にある農家属性を明らかにする。

1つめの親ノード「基幹的農業者の平均年齢」に対する確率推論を行った結果を図6に示す。グラフの縦軸は推論によって求められた事後確率を表す。

島原地域において確率推論を実行すると、基幹的農業従事者の平均年齢が「70代」を境に営農継続と離農の確率値は対称的に推移した。平均年齢の若い「50代以下」「60代」で営農継続農家の確率値が離農農家を12.6～16.1%上回り、平均年齢が「80代以上」と「基幹的従事者がいない」では離農農家が営農継続農家を14.8～16.7%上回った。このことは、基幹的農業従事者の平均年齢が60代以下の経営体は営農を継続する可能性が高く、平均年齢が80代以上の高齢農家や基幹的農業従事者不在の経営体では離農の可能性が高まることを示している。

県北地域では「80代以上」で離農農家が営農継続を

大きく上回り、この地域でも高齢農家による離農の傾向が読み取れた。ただ、島原地域とは異なり、基幹的農業従事者の平均年齢が「60代以下」でも営農継続農家と離農農家との確率の差は小さく、また、「基幹的農業従事者がいない」では両者に差がみられなかった。つまり、基幹的農業従事者の平均年齢が70代以下であれば年代の差による影響は少なく、また、基幹的農業従事者の有無が営農継続/離農に対して与える影響はほとんどないと解釈できる。

同様の方法により得た「経営耕地面積」に対する確率推論結果を図7に示す。

島原地域では階級の境界値である耕地面積100aを境に、営農継続農家と離農農家の確率値は対称的に推移した。営農継続農家では経営規模が100aより大きいカテゴリほど確率値が増加し、「150a以上」と回答した農家では営農継続農家が離農農家を20.0%も上回った。逆に「50a未満」では、離農農家が営農継続農家を22.0%上回った。島原地域では、経営耕地面積の大小が営農継続/離農に非常に大きな影響を与えていることがわかる。

県北地域は「50a未満」で離農農家が営農継続農家を20.1%上回り、耕地面積50a未満の零細規模の経営体で顕著な離農の傾向がみられた。一方、経営規模50a以上の各カテゴリでは営農継続が離農を上回るものの、その差は6.0～6.5%程度と小さく、50a以上規模の大小による傾向の違いは見られなかった。「農産物販売額」に対する確率推論結果を図8に示す。

島原地域では、「50万～300万円」のカテゴリを境に営農継続と離農が対称的に推移した。販売金額「300万～1000万円」および「1000万円以上」で営農継続農家が離農農家を大きく上回った。一方で、「50万円未満」では営農継続農家の確率値は大幅に低下し、「販売なし」では1.7%にまで低下した。反対に、離農農家においては「50万円未満」のカテゴリで営農継続農家の確率値を大きく上回ることから、販売金額50万円が営農継続/離農の大きな転換点となっていると考えられる。

県北地域は「販売なし」で営農継続農家と離農農家との差が大きくなり、農産物販売額の有無が影響していた。だが、これ以外のカテゴリでは営農継続と離農との確率値の差は小さく、販売金額の規模による傾向の違いはみられなかった。

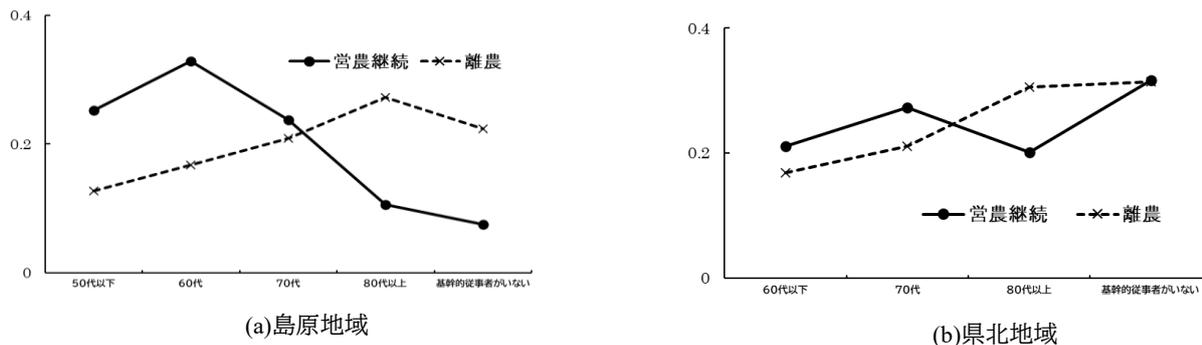


図6 「営農継続/離農」に観測値を与えたときの「基幹的農業従事者の平均年齢」の推論結果

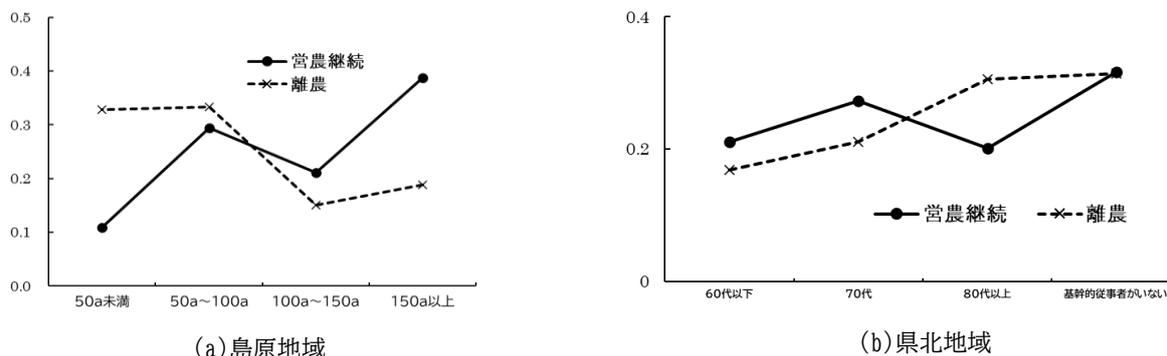


図7 「営農継続/離農」に観測値を与えたときの「経営耕地面積」の推論結果

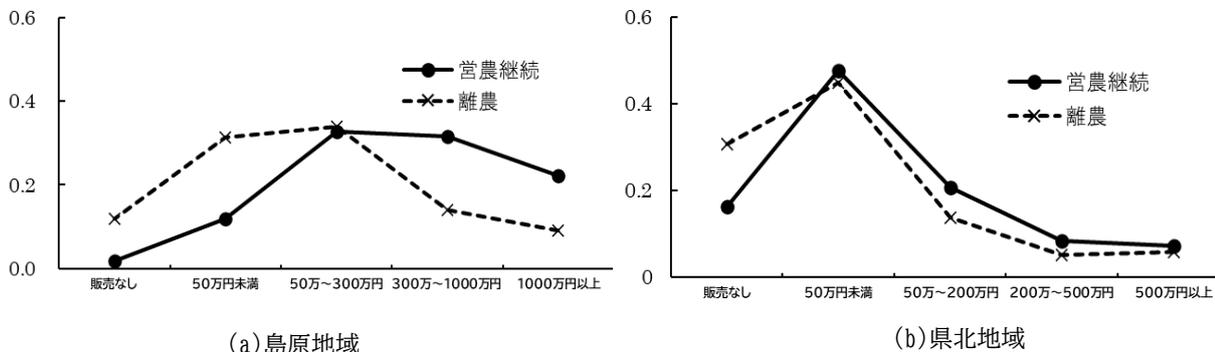


図8 「営農継続/離農」に観測値を与えたときの「農産物販売額」の推論結果

### 3) 単純集計と確率推論結果に基づく解釈

土壌条件に恵まれた島原地域は、専業農家を中心にした安定した経営基盤を有する経営体が多く存在し、後継者への事業継承も順調に行われている地域である。確率推論の結果から、この地域の営農継続農家の特徴として、基幹的農業従事者の平均年齢が60代以下であること、経営耕地面積100a以上であること、農産物販売額300万円以上であることが挙げられる。安定した生産基盤を有する担い手を確保することにより、今後も持続的な農業の発展が見込まれる地域である。ただ一方で、この地域の離

農農家の特徴である基幹的農業従事者不在経営体の割合は近年増加しており、潜在する離農意向農家の把握に努め、離農農家が所有する耕作条件の良い農地を意欲ある担い手へと流動化を進めていく必要がある。

他方、営農条件に様々な制約のある県北地域は、基幹的農業従事者不在の経営体がおよそ半数を占め、農業を副業的に行っている零細規模の農家も多く存在する地域である。確率推論によってこの地域の営農継続農家の傾向を明らかにすることはできなかったが、離農農家の特徴については読み取るこ

とができた。その特徴としては、基幹的農業従事者の平均年齢が 80 代以上の高齢農家であること、経営耕地面積が 50a 未満であること、農産物販売額がないこと、が挙げられる。単純集計結果から、農産物販売額「50 万円～300 万円未満」の経営体数の割合が 5 年間で 6% 近く増加していることは、地域農業の維持に向けてポジティブな材料と言える。ただし、経営耕地面積が 50a 未満の経営体が地域全体の

およそ 2 割を占め、世代交代が進まないまま高齢化が進展していることを考慮すれば、集落によっては今後、農地の保全・管理だけでなく農村地域コミュニティの弱体化が懸念される。経営規模や専業・副業に関わらず、多様な人材を「地域の担い手」として確保し、地域の多面的機能の維持を図っていく必要がある。

## 5. まとめ

本研究では、センサス個票データを用いてベイジアンネットワークモデルを構築し、本県 2 地域の農業動向分析を試みた。まずは単純集計結果から、地域農業を支える経営体の特徴の差異やその変化を捉えた。そのうえで、構築したモデルを用いて営農継続または離農に対して直接的な影響を与えている親ノードに対する確率推論を実施した。単純集計結果を併せて推論による確率値を解釈することにより、今後のリスクの予見とその対策について検討した。

ベイジアンネットワークは各事象間にある因果関係を視覚的に表現することで、既知の知見と照らし合わせながら、人間の意思決定に近いプロセスでモデルを構築することができる。また、確率変数間

の関係性を定量化する確率推論は、原因から結果だけでなく、結果から原因を予測することもできる柔軟なモデルである。これまでにベイジアンネットワークによる離農予測はイタリアでの事例 (Gambelli and Bruschi, 2010) が報告されているが、日本における研究は見当たらなかった。中山間地域を多く抱える本県では、2045 年の総人口指数が全国で 8 番目に低い 71.3 と推計されており、このときの老年人口割合は 40.6% に達するとされる (国立社会保障・人口問題研究所, 2018)。この大きく変動する不確かな状況下において、今後想定されるリスクとこれに対する政策の効果を確率値として評価できるベイジアンネットワークは、政策立案者にとって有用な意思決定支援ツールになり得ると考えられた。

## 6. 摘要

長崎県の農業構造において少子高齢化の進展は深刻な問題であり、今後農業生産基盤の脆弱化が懸念される。本研究では、本県の島原地域と県北地域の 2 地域の耕種農家の農業センサス個票データを用いてベイジアンネットワークによる農業動向分析を行った。2015 年～2020 年の 5 年間に営農継続または離農した農家分類を目的変数とし、前処理を施したセンサス個票データの調査項目を説明変数としてベイジアンネットワークモデルを構築し、農業構造において複雑に絡む複数の事象間の因果関係

を視覚的に表現した。また、目的変数に直接的な影響を与えている親ノードに対する確率推論を実行することにより、それぞれの地域の営農継続または離農しやすい経営体の特徴を捉えることができた。

ベイジアンネットワークによる農業動向分析は、今後のリスクの予見と、これに対する政策の効果を確率値として評価できることから、政策立案者にとって有用な意思決定支援ツールになり得ると考えられた。

## 7. 引用文献

- Gambelli, D. and V. Bruschi. 2010. A Bayesian network to predict the probability of organic farm's exit from Marche, Italy. *Computers and Electronics in Agriculture*. 71: 22-31.
- 細山高夫. 2012. 北海道における農業構造変化の地域特性と将来動向, 北海道農業研究センター農業経営研究. 106: 1-47
- 国立社会保障・人口問題研究所. 2018. 日本の地域別将来推計人口(平成30(2018)年推計). <https://www.ipss.go.jp/pp-shicyoson/j/shicyoson18/3kekka/Municipalities.asp>. 2023年6月30日参照
- 栗原伸一・霜浦森平・丸山敦史. 2014. 農業センサ個票データを用いた離農要因の地域別分析. *農業情報研究*. 23(2): 72-81
- 本村陽一・佐藤泰介. 2000. ベイジアンネットワークー不確定性のモデリング技術ー. *人工知能学会誌*. 15(4): 575-582
- 本村陽一. 2003. ベイジアンネットワークによる確率的推論技術. *計測と制御*. 42(8): 649-654
- 農林水産省. 2015. 2015年農林業センサ農業経営体調査票. <https://www.e-stat.go.jp/stat-search/fi>le-download?statInfId=000031510318&fileKind=2. 2023年6月30日参照
- 農林水産省. 2022a. 令和4年耕地面積(7月15日現在). <https://www.maff.go.jp/j/tokei/kouhyou/sakumotu/menseki/index.html#r>. 2023年6月30日参照
- 農林水産省. 2022b. 令和3年農業総産出額及び生産農業所得(都道府県別). [https://www.maff.go.jp/j/tokei/kouhyou/nougyou\\_sansyutu/index.html#r](https://www.maff.go.jp/j/tokei/kouhyou/nougyou_sansyutu/index.html#r). 2023年6月30日参照
- 大竹文雄・内山 融・小林庸平. 2022. エビデンスに基づく政策形成の導入と実践. 日本経済新聞出版
- 繫枺算男・植野真臣・本村陽一. 2006. ベイジアンネットワーク概説. 株式会社培風館
- 寺谷 諒. 2018. 機械学習を用いた地域農業動向予測モデルの構築に関する研究. 京都大学情報学研究科博士論文
- 安武正史. 2018. マルコフモデル利用による予測結果の検証と2015年農林業センサデータの有効性. *関東東海北陸農業経営研究*. 108: 53-57

## Summary

The declining birthrate and aging population are serious problems in the agricultural structure of Nagasaki Prefecture, and there are concerns that the agricultural production base will become more vulnerable in the future.

This study uses the data of individual farmers in the Shimabara and Kenhoku regions of Nagasaki Prefecture to analyze the agricultural census data. In this study, I analyzed agricultural trends using a Bayesian network model based on individual farmer census data in two regions of Nagasaki Prefecture, Shimabara and Kenhoku, using the classification of farmers who continued or quit farming during the five-year period from 2015 to 2020 as the objective variable and survey items from the pre-processed individual census data as explanatory variables. A Bayesian network model was constructed to visually represent the causal relationships among multiple events that are intricately intertwined in the agricultural structure. Probabilistic inference was performed on the parent nodes that have a direct effect on the target variable. As a result, the attributes of continued farming and exited farming in each region were captured.

The analysis of agricultural trends using Bayesian networks allows for the forecasting of future risks and the assessment of policy effects in terms of probability values. This makes it a potentially valuable decision-support tool for policymakers.