

都道府県別の年収に関する統計的分析

2010SE221 高井翔平

指導教員：松田眞一

1 はじめに

今年、私が就職活動の年ということもあり、自分のなりたい職業の年収について調べていった。そこで、都道府県別ではどのような違いがあるのか疑問を持った。平成 22 年に国勢調査が行われ、産業別就業者に関するデータが都道府県別に明確にされた、現在の日本の現状はどのようなものなのかを分析してみようと思う。

2 データについて

平成 20 年に行われた、厚生労働省による賃金構造基本統計調査から各都道府県の平均年収 ([1] 参照)、総務省統計局平成 22 年国勢調査の統計表第 16 章の 5 に、19 種類に業種分類された都道府県別の就業者数のデータ ([2] 参照)、内閣府平成 19 年の統計情報による都道府県別の全国物価地域差指数 ([3] 参照) を用いた。

3 分析方法

使用した分析方法は重回帰分析、クラスター分析、主成分分析である。(金 [4] 荒木 [5] 小西 [6] 永田・棟近 [7] 参照)

4 重回帰分析

step による変数選択の際に残った 12 個の変数を VIF(分散拡大要因) で調べた結果、漁業、製造業、情報通信業、複合サービス事業、物価地域差指数の 5 個の変数が残った。残差分析をした結果、正規性があり、外れ値もなかった。解析したところ、決定係数が 0.6649、修正済み決定係数が 0.622 となった。

表 1 step による変数選択をした後に VIF で調べた解析結果

| | 推定値 | 標準誤差 | t 値 | P 値 |
|-----|-----------|----------|--------|--------|
| 定数項 | -142.2820 | 283.9095 | -0.501 | 0.6191 |
| 漁業 | -2.5135 | 1.9481 | -1.290 | 0.2046 |
| 製造 | 0.2543 | 0.0605 | 4.204 | 0.0001 |
| 情報 | 0.0393 | 0.1684 | 0.233 | 0.8166 |
| 複合 | -6.2986 | 3.1903 | -1.974 | 0.0555 |
| 物価 | 5.9273 | 2.9190 | 2.031 | 0.0492 |

製造業が最も影響が強い変数となっている。それに続き、物価地域差指数、複合サービス事業となった。製造業と物価地域差指数が正の影響を、漁業と複合サービス事業が負の影響を及ぼす変数と考えられる。

製造業は「ものづくりの国」と日本が呼ばれるだけあって値が高いのではないかと推測できる。また、日本の輸出品の上位が自動車や自動車の部品、半導体等の電子部品など

の高い技術を有する物が占めていることも原因の一つだと考えられる。

複合サービス事業とは郵便局、農業協同組合等が分類される職業である。郵便局、農業協同組合という、都道府県の大きさによっても違うが、基本的にあまり変わらない傾向にある。よって、複合サービス事業の割合が多いということは全体的に就業者が少ないことである。就業者が少ないことは他県へ出稼ぎに行く人々が多いことが推測されるため、平均年収に関係していると考えられる。

5 クラスタ分析

本研究では、安定性があったためウォード法を使用した。職種の違いや物価の影響などが見たかったため、データを標準化をし、デンドログラム図 1 を出力した。ここで言う標準化とは、各都道府県の就業者数の全体を 1 として、百分率でそれぞれの業種の就業者を表した。大きく分けて左から 5 群に分けることにした。

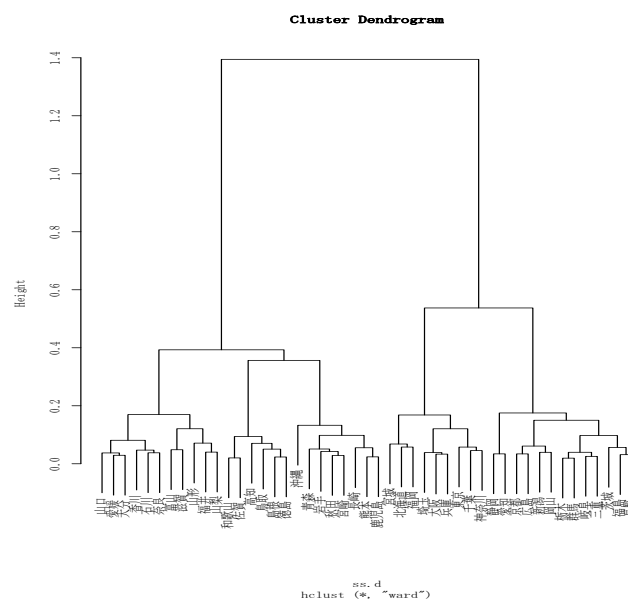


図 1 クラスタ分析によるデンドログラム

5.1 第 1 群

第二次産業である製造業、建設業、電気関係が多い。労働人口、平均年収、物価指数は、それぞれ平均より下である。第二次産業に関わる労働者の割合が多いが、労働人口がやや少ない群である。

5.2 第 2 群

第一次産業である農業や漁業が盛んな群である。この群は就業者が最も少ない群ということもあり、複合サービス

や医療関係、公務の割合が多いと考えられる。また、就業者が一番少ない群であるが、平均年収と物価指数は最も低いわけではないことが分かった。

5.3 第3群

農業、漁業をする就業者が最も多い群である。観光地が多いこともあり、生活関連サービス業、娯楽業をする就業者の割合が多いことが推測できる。よって第一次産業、第三次産業が盛んであり、最も平均年収と物価指数が低い群である。

5.4 第4群

労働人口、平均年収、物価指数、それぞれ最も高い群である。情報通信業、運輸業、卸売業、金融業、不動産業などの第三次産業の割合が他の群と比べて明らかに多い。様々な職種の就業者の割合が多いことにより、複合サービスや医療関係、公務などの割合が少ない。都会、または都会に隣接している県の群であり、様々な専門的な職種が集まっている傾向がある。

5.5 第5群

主に製造業、電気・ガス・熱供給・水道業などの第二次産業の割合が多く、次に農業などの第一次産業が多く、第4群の次に労働人口、平均年収、物価指数が高い群である。他の群と比べて、製造業以外の職種の割合は特徴的なものがあまりなく、製造業に特化している都道府県と推測できる。

6 主成分分析

重回帰分析の際に変数選択したものから物価地域差指数を引いた4つの変数を使い相関行列に基づく主成分分析をしたものを示す。主成分の数の決め方は本研究において、累積寄与率70%以上を適用することにした。

表2 固有値、寄与率、累積寄与率

| | 第1主成分 | 第2主成分 | 第3主成分 |
|-------|--------|--------|--------|
| 固有値 | 1.4670 | 1.2827 | 0.4502 |
| 寄与率 | 0.5380 | 0.4113 | 0.0507 |
| 累積寄与率 | 0.5380 | 0.9493 | 1.0000 |

表3 主成分得点

| | 第1主成分 | 第2主成分 |
|----------|--------|--------|
| 漁業 | 0.589 | 0.308 |
| 製造業 | -0.572 | 0.424 |
| 情報通信業 | 0.224 | -0.735 |
| 複合サービス事業 | 0.526 | 0.429 |

6.1 第1主成分

正の値になっているものが漁業、情報通信業、複合サービス事業であり、第一次産業、第三次産業の要素で構成されている。負の値になっているものが製造業となり、第二次産業の要素で構成される。よってこの軸は、第一次産業および第三次産業と第二次産業の比較の軸である。負の値にいくにつれ、平均年収が高くなっている傾向にある。

6.2 第2主成分

正の値になっているものが漁業、製造業、複合サービス業である。負の値になっているものが情報通信業となる。よってこの軸は、情報通信業の職種が多い都道府県かどうかの軸である。こちらも負の値にいくにつれ、平均年収が高くなっている傾向になっている。

7 まとめ

都道府県別の平均年収に関係するものの大きな要因は、製造業の割合が大きく影響することが分かった。製造業は国内で最も就業者が多い職業の一つである。特に第二次産業などの専門的な職業が盛んな地域では就業者数、平均年収、平均物価指数のそれぞれが高い数値となっている。

もう一つの傾向としては、生活に欠かせない職業である複合サービスや医療関係、公務などの割合も大きく平均年収に関係することである。これらの割合が多いということは、他の様々な専門的な職業に就いている就業者が少なく、その地域の就業者も少ないことが考えられる。

8 おわりに

本研究では平均年収に関して解析を行なっていく、各都道府県の特徴など、より深い知識が得られた。現在の日本では、職業による就業者の割合により平均年収の違いが顕著に見られた。

参考文献

- [1] 都道府県年収ランキング <http://nensyu-labo.com/>
- [2] 総務省統計局統計表 16-5 都道府県、産業別就業者数 (平成22年) <http://www.stat.go.jp/data/nihon/16.htm>
- [3] 全国物価地域差指数ランキング (都道府県別) <http://www.keizai-bank.com/ranking/todofuken-bukka.html>
- [4] 金明哲:『Rによるデータサイエンス:データ解析の基礎から最新手法まで』。森北出版社, 2007.
- [5] 荒木孝治:『RとRコマンドーではじまる多変量解析』。日科技連出版社, 2007.
- [6] 小西貞則:『多変量解析入門—線形から非線形へ』。岩波書店, 2010.
- [7] 永田靖・棟近雅彦:『多変量解析入門』。サイエンス社, 2001.