

自殺に関する統計的分析

2005MM061 佐藤のぞみ

指導教員：木村美善

1 はじめに

最近では、いじめ、不登校、殺人、子供が親を殺すなど考えられないニュースを耳にする。今の人たちはどんな風に考え生きているのだろうか。その中でも特に、自分の命を絶つほど苦しんでいる人の気持ちを少しでも理解したいと思ってこのテーマに決めた。毎年日本では3万人以上の方が自殺する。毎日およそ80人が自殺するという計算になる。何が原因で自殺するのだろうか。自殺するほど苦しい状況に追い込ませているのは一体何なのか。経済的な問題、気候条件、社会制度など、影響を与えていると考えられる問題から自殺にまで追い込ませているものを調べていきたいと思う。

2 データについて

ホームページから、以下のデータを入手した。

- (1) 原因・動機別・年代別自殺者数 (平成19年)
- (2) 原因・動機別・職業別自殺者数 (平成19年)
- (3) 年代別自殺者数の推移 (昭和53年～平成19年)
- (4) 都道府県別自殺動機自殺数 (平成16年～平成18年)
- (5) 都道府県・職業別自殺数 (平成16年～平成18年)
- (6) 自殺率とそれを説明する変数のデータ (平成13年～平成18年)

3 分析方法

分析方法として、クラスター分析法、主成分分析法、重回帰分析法を用いた ([1],[2] 参照)。

4 自殺率の特徴

4.1 クラスター分析

県により自殺の状況に違いがあるのか考察するためにクラスター分析した。まず、職業により県を分類することを考え、平成16年から18年の都道府県・職業別自殺率のデータ (縦に47都道府県、横に職業5項目) を用いた。次に、都道府県・自殺動機別自殺率のデータ (縦に47都道府県、横に自殺動機8項目) を用いて、同様な分析をした。どちらも基準化を行って分析を行う。ここでは、前者の分析結果のみ載せておく ([3],[5] 参照)。

「第一群」:自殺率が割と低く、特に学生の自殺率が低い群、「第二群」:自殺率の高い群、「第三群」:自殺率が割と高く、特に学生・被雇用者の自殺率が高い群、「第四群」:自殺率の低い群

このように、自殺率の高さによって群分けされた。東北地方のような第二群に属する群は、気温が低く高齢者の割合が多く、医療資源が不十分なため自殺率が高くなるからだと考えられる。また、逆に第四群の都会の地方に位置する自殺率の低い県は、社会環境が整っており、高齢化社会

と言われる現在のための医療設備も整っているからだと考えられる。このように、県による自殺率の高低が現れ、県による群分けをすることが出来た。

4.2 主成分分析

都道府県・職業別自殺率のデータを基準化して主成分分析を行った。すると、第二主成分までの累積寄与率が79%となった。第一主成分と第二主成分はそれぞれ以下のような式になる。

$$Z_1 = -0.53x_1 - 0.54x_2 - 0.56x_3 - 0x_4 - 0.33x_5$$

$$Z_2 = 0x_1 - 0.22x_2 + 0x_3 - 0.92x_4 + 0.31x_5$$

第一主成分は、学生を除いてすべての係数が負となる。これは、学生以外の自殺率の高さをみる軸であることが分かる。また第二主成分は、管理職が正で、被雇用者・学生が負となる。これは、管理職での自殺率の高さ、被雇用者・学生の自殺率の高さをみる軸であることが分かる。

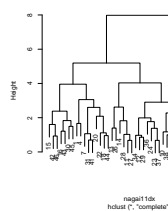


図1 職業別自殺率

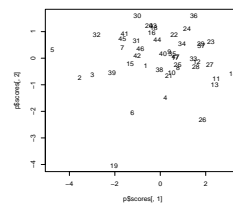


図2 主成分得点

5 自殺の動機の特徴

自殺動機が年齢とどのように関係しているのかを調べるためにクラスター分析を用いた。まず、年齢により自殺動機を分類することを考え、原因・年代別自殺率 (縦に原因・動機7項目、横に年齢6項目) を用いた。次に、自殺動機の詳細データ (縦に原因・動機52項目、横に年齢6項目) を用いて同様な分析をした。ここでは、後者の分析結果のみ載せる ([4] 参照)。

「第一群」:2群から4群以外の理由での自殺、「第二群」:身体疾患での自殺、「第三群」:うつ病での自殺、「第四群」:負債・統合失調症での自殺

52項目を四群に群分けした結果、第一群は46項目、第四群は4項目、第二、三群は1項目選ばれた。第一群は、自殺率が低い動機の群が集まっていて、第二、三、四群は自殺率が高く中でも第三群のうつ病による自殺率が高い。

6 自殺者の職業の特徴

自殺する人はどういう人なのか、職業が自殺動機とどのように関係しているのかを調べるためにクラスター分析

をした。まず、自殺動機により職業を分類する事を考える、原因・職業別自殺者数(縦に職業5項目、横に自殺動機7項目)を用いた。次に、自殺動機の詳細データ(縦に職業30項目、横に自殺動機7項目)を用いて同様な分析をした。ここでは、後者の分析結果のみ載せる([4]参照)。

「第一群」:主婦・年金雇用者の自殺,「第二群」:失業者・管理職・サービス業の自殺,「第三群」:勤め人・事務職・自営業などの自殺,「第四群」:中高生の自殺

30項目を四群に群分けした結果、第三群は27項目、第一、二、四群は1項目に群分けされた。一群は家庭・健康問題での自殺が多く、二群は経済・勤務・男女問題での自殺が多い。また、四群は学校・男女問題での自殺が多い。

7 自殺率の回帰分析

ここでは、県別データを用いて、自殺率がどのような要因によってどのように説明されるかを回帰分析によって調べてみる。まず、平成13年度から平成18年度までのデータを用いる。各年度ごとに、自殺率を説明する要因から自殺率がどのように説明できるのか、重回帰分析を行う。また、重回帰分析を行う際に、変数選択法(ステップワイズ法)を用いて変数選択を行う([5]参照)。

7.1 平成13年度から18年度までの6年間の分析

6年間分を一つのデータとして重回帰分析してみることにする。説明変数は、 $x_1=65$ 歳以上の割合、 $x_2=離婚率$ 、 $x_3=平均気温$ 、 $x_4=日照時間$ 、 $x_5=給与額$ 、 $x_6=労働者の割合$ 、 $x_7=病院施設数割合$ 、 $x_8=病床数割合$ 、 $x_9=大学・公的以外の医療費の割合$ 、 $x_{10}=失業率$ の10変数である。重回帰分析をして、変数選択を行った結果以下のような結果が得られた。重回帰式は $y=16.81+1.2x_1-0.3787x_3-0.0028x_4-0.0181x_5-23.13x_7+0.2981x_8+0.0009x_9+1.714x_{10}$ で、決定係数が0.6413、修正済決定係数が0.6307となった。 x_1, x_3, x_9, x_{10} のp値がかなり小さいので、65歳以上の割合、平均気温、医療費の割合、失業率がよく説明されていることが分かる。そこで、選ばれた変数 $x_1, x_3, x_4, x_5, x_7, x_8, x_9, x_{10}$ を用いて各年度ごとに重回帰分析を行う。以下、決定係数を R^2 、自由度修正済み決定係数を R^{2*} とし、結果を述べる。

平成13年 $y=20.6851+1.0377x_1-0.2619x_3-0.0027x_4-0.0226x_5-9.7498x_7-0.2186x_8-0.0003x_9+1.5039x_{10}$ で、 $R^2=0.4993$ 、 $R^{2*}=0.3939$ 。

平成14年 $y=28.47+1.194x_1-0.324x_3-0.0068x_4-0.0251x_5-8.4279x_7+0.3167x_8-0.0013x_9+1.7469x_{10}$ で、 $R^2=0.7155$ 、 $R^{2*}=0.6556$ 。

平成15年 $y=42.874+0.836x_1-0.4206x_3-0.0053x_4-0.0834x_5-31.6303x_7+0.351x_8-0.0007x_9+1.394x_{10}$ で、 $R^2=0.6251$ 、 $R^{2*}=0.5462$ 。

平成16年 $y=28.7844+1.2384x_1-0.2559x_3-0.0071x_4-0.0193x_5+32.6155x_7+0.1317x_8-0.0015x_9+1.8931x_{10}$ で、 $R^2=0.757$ 、 $R^{2*}=0.7058$ 。

平成17年 $y=30.1639+0.9235x_1-0.5339x_3-0.0038x_4-0.0211x_5-3.0331x_7+0.558x_8-0.0014x_9+1.5737x_{10}$ で、 $R^2=0.7615$ 、 $R^{2*}=0.7113$ 。

平成18年 $y=26.4576+1.0741x_1-0.2488x_3-0.0011x_4-0.0485x_5-45.428x_7+0.8587x_8-0.0015x_9+1.5803x_{10}$ で、 $R^2=0.7301$ 、 $R^{2*}=0.6733$ 。

7.2 考察

各年度でp値が0に近いのは、 x_1 と x_{10} のp値である。つまり、65歳以上の割合と失業率が、自殺率を説明するのに大きな影響を与えていることが分かる。このことから、高齢者の割合と失業率が増えれば増えるほど、自殺率も高くなることが言える。また、各年度の x_7 のp値をみると、0.05より大きな値をとっているの、よく説明出来ていないことが分かる。その他の、 x_3, x_4, x_5, x_8, x_9 の係数を見ると、年度毎の違いはほとんどみられない。 x_3, x_4, x_5, x_9 の係数は、マイナスの値をとり、 x_8 は平成13年を除いてプラスの値をとっている。このことから、平均気温や日照時間、給与額、医療費が増えれば、自殺率が下がり、病床数の割合が増えれば自殺率が上がることが分かる。これは、病床数の増加が直接自殺率の増加に結びついているということではなく、高齢者の増加に伴って病床数を増やしたところで年々増加していく自殺率をくいとめることは出来ないということを示唆していると考えられる。ただし、 x_3, x_4, x_5, x_8, x_9 のp値が0.05より大きい値の場合が多いので、あまり説明されているとは言えない。

8 まとめ

この研究を通して、自殺率に最も影響を与えている要因として、高齢者の割合と失業率が考えられる。その他にも、平均気温、給与額、医療費の割合が高くなれば、自殺率は低くなることが分かった。自殺を防ぐために、失業、倒産、長時間労働が起こっている現在、社会制度の見直しや相談・支援体制の整備といった社会的な取り組みが必要である。また自殺をする人は、さまざまな悩みにより心理的に追い詰められ、自殺以外の選択肢が考えられない状況に陥っている。私たちに出来ることとして、周りの人であつ病になっている人、またはうつ病になりかけそうな人がいたら、親身になって悩みを聞いてあげたり、専門家への相談やうつ病などの治療を勧めることが一番である。

参考文献

- [1] 間瀬茂・神保雅一・鎌倉稔成・金藤浩司: 工学のためのデータサイエンス入門, 数理工学社, 2004.
- [2] 白旗慎吾: 統計解析入門, 共立出版, 1992.
- [3] 自殺実態白書
<http://www.lifelink.or.jp/hp/whitepaper.html>
- [4] 警視庁: 平成19年度の自殺の概要資料
http://www.npa.go.jp/toukei/chiiki10/h19_zisatsu.pdf
- [5] 統計局・政策統括官・統計研修所
<http://www.stat.go.jp/index.htm>