

資料2 分析の手法等

1 統計的手法による検証（第1章、第2章関係）

みえ県民意識調査は、一部の標本を抽出し、その結果から全体の値を推定する「標本調査」です。

この調査では5,710の回答数（サンプル数）がありますが、調査結果と県全体の本当の姿との間にはどうしても誤差（＝標本誤差）が発生します。また、属性を組み合わせで細分化すると、そのカテゴリーのサンプル数はさらに少なくなることから、誤差はより一層拡大します。

このため、幸福感の平均値や暮らしの実感の比率に差があったとしても、結果として、そのことがそのまま県民全体に当てはまるとは言い切れない（統計的に有意ではない）ケースが考えられます。

そこで、幸福感の平均値や暮らしの実感の比率について比較を行うにあたり、その差に統計的な有意性があるかどうか、ここでは、同じ調査を異なる調査対象で100回行った場合、95回以上の割合で同様の差が生じると言えるかという観点から、下記の検定方法により判定を行いました（1、2）。

1 幸福感の平均値の差の検定方法

$$U = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{S_1^2}{n} + \frac{S_2^2}{m}}}$$

\bar{X} : 標本Xの平均値 \bar{Y} : 標本Yの平均値
 S_1^2 : 標本Xの分散 S_2^2 : 標本Yの分散
 n : Xのサンプル数 m : Yのサンプル数

$U > 1.64$ の時、平均値の差は統計的に有意であると言える(危険率5%)

2 暮らしの実感の比率の差の検定方法

$$U = \frac{P_1 - P_2}{\sqrt{\frac{P_1 \times (1 - P_1)}{n} + \frac{P_2 \times (1 - P_2)}{m}}}$$

P_1 : 標本Xの回答比率 P_2 : 標本Yの回答比率
 n : Xのサンプル数 m : Yのサンプル数

$U > 1.64$ の時、回答比率の差は統計的に有意であると言える(危険率5%)

2 分析の手順（第2章関係）

(1) クロス集計

問2の各設問と各属性でクロス集計を行いました。

なお、「不明」及び「わからない」回答については除いて分析を実施しています。

【例】問2 - 1（災害等の危機への備えが進んでいる）と問7 - 1（性別）の「クロス集計」

		2 - 1災害等の危機への備えが進んでいる				総計
		感じる A	どちらかといえば感じる B	どちらかといえば感じない C	感じない D	
7-1 男性		75	538	954	929	2496
女性		114	645	1058	794	2611
男女計		189	1183	2012	1723	5107

(2) 回答比率の算定

クロス集計結果から、属性項目毎に回答項目の比率を算定しました。

【例】問2 - 1（災害等の危機への備えが進んでいる）と問7 - 1（性別）の「回答比率」

		2 - 1災害等の危機への備えが進んでいる				総計
		感じる	どちらかといえば感じる	どちらかといえば感じない	感じない	
7-1 男性	÷E	3.0%	21.6%	38.2%	37.2%	100.0%
女性	÷E	4.4%	24.7%	40.5%	30.4%	100.0%
男女計	÷E	3.7%	23.2%	39.4%	33.7%	100.0%

(3) 比率の差の算定

それぞれの属性項目と属性全体（＝属性項目の合計）の回答比率の差を算定しました。

【例】問2 - 1（災害等の危機への備えが進んでいる）と問7 - 1（性別）の「比率の差」

		2 - 1災害等の危機への備えが進んでいる				総計
		感じる	どちらかといえば感じる	どちらかといえば感じない	感じない	
7-1 男性	' - '	-0.7%	-1.6%	-1.2%	3.5%	
女性	' - '	0.7%	1.5%	1.1%	-3.3%	
男女計						

(4) 差の検定

回答比率の差が統計的に有意な差があるか検定を行いました。

検定結果(U)の絶対値が1.64を超える時、比率の差は統計的に有意であると言えます(危険率5%)。検定方法の詳細については前項の「統計的手法による検証」を参照

【例】問2 - 1（災害等の危機への備えが進んでいる）と問7 - 1（性別）の「差の検定」

		2 - 1災害等の危機への備えが進んでいる				総計
		感じる	どちらかといえば感じる	どちらかといえば感じない	感じない	
7-1 男性	U	-0.290	-0.747	-0.614	1.783	
女性	U	0.282	0.735	0.604	-1.672	
男女計						

上記の例では、男性は全体よりも「感じない」と回答した人の比率が高く、女性は全体よりも「感じない」と回答した割合が低くなっており、またこれらの差は、いずれも統計的に有意であると言えます。

このことから、「『災害等の危機への備えが進んでいる』と感じない人は女性よりも男性に多い」と言えます。

3 分析の手法等（第3章関係）

（1）基本的考え方

以下は、第3章の分析にあたって前提となる基本的な考え方です。

相関係数については、ピアソンの積率相関係数を算定することとします。

相関係数はいくつかの種類があり、対象となるデータにより使い分ける必要があります。今回の分析データについて、問1の回答データ(0～10点)は間隔尺度であり、また、問2の回答データ(「感じる」、「どちらかといえば感じる」、「どちらかといえば感じない」、「感じない」)は順序尺度ですが、調査票の選択肢に1～4の数字を振ってあることから、今回の分析において間隔尺度とみなしても差し支えないと考えられます。

このように判断し、問1と問2の回答間の関係を分析するにあたり、間隔尺度以上のデータに適用できる相関係数の算定及び回帰分析を行うこととしたものです。

問1と問3の関係分析、問2及び問3の設問間の関係分析についても、同様です。

〔ピアソンの積率相関係数 = XとYの共分散 ÷ (Xの標準偏差 × Yの標準偏差)〕

なお、算定の際、問2、問3の回答については、調査票の選択肢の番号「1、2、3、4」を、それぞれ「4、3、2、1」という数字に置き換え計算します。例えば、問2では、「感じる」を「4」、「どちらかといえば感じる」を「3」、「どちらかといえば感じない」を「2」、「感じない」を「1」とします。

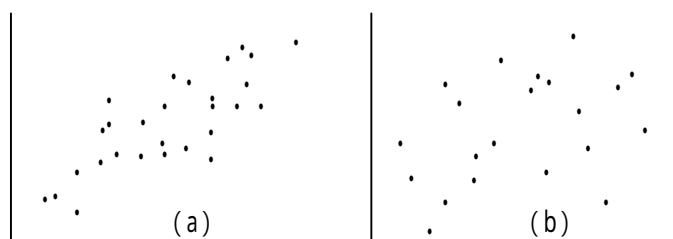
個々人の「幸福感」がさまざまな要因に影響を受けると考えられる中で、今回の分析においては、生活の一部の側面に係る実感と「幸福感」との関係を対象としていることから、その関係を表す相関係数や回帰分析における決定係数は必然的に高い値になりにくくなります。このため、相関係数等の値が低くても(どれくらいであれば低いのかは一概に言えませんが、例えば0.1や0.2であっても)、統計的に有意(母集団の相関係数 0、回帰係数 0とみなせる)であれば、その相関係数等は十分意味を持つものとして分析を行います。

一方、相関係数等の値が低い場合は、相関係数そのものは有意であっても、相関係数の差が有意とならない可能性が高いことから、本章では、母集団(県民全体)に関する推定は行わず、標本(有効回答のあった5710人)について言える事に絞り、関係の把握に努めることとします。

（2）相関係数

相関係数は、二つの変数(いろいろな値をとりうるもの)の直線的な関係の強さを表す指標で、-1から+1までの値をとります。一方の変数の値が大きいほど他方の変数の値も大きくなる傾向があるとき、「相関係数」は正となり、正の相関関係があると言います。相関係数が1(或いは-1)に近いほど、2つの変数の分布を描いた散布図上の点が、ある直線の周りに集中する、つまり、相関関係が強いということになります。

左の散布図(a)、(b)は、どちらも正の相関関係がある場合のイメージです。相関係数の値は、(a) > (b) となります。



【相関関係と相関係数】

2つの変数 x 、 y の間で、一方が増加すると他方が増加または減少するような関係を相関または相関関係といいます。相関関係は、以下の3つの場合に分けることができます。

正の相関: x が増加(減少)すると y も増加(減少)する。

負の相関: x が増加(減少)すると y が減少(増加)する。

無相関: x の増加・減少が、 y の増加・減少と直線的な対応をしない。

相関関係の強さを数値で表したものが相関係数で、必ず -1 から1 の間の値をとります。

また、相関係数 r の値は、おおよそ次のように評価することができます。

r が1に近いほど強い正の相関関係がある。

r が -1 に近いほど強い負の相関関係がある。

r が0に近いほど相関関係は弱い。

(「よくわかる統計学 基礎編[第2版](2011年、金子治平・上藤一郎編)」より)

【相関係数の読み方】

相関係数の読み方については、次のような考え方もあり、専門家のこうした見解を考慮に入れ、分析を行う必要があります。

どのくらいの相関関係の値があれば「高い相関がある」といえるかということがしばしば問題になる。…(中略)…相関係数の値が大きい小さいかは、あくまでも相対的なものであって、いくらであれば「高い相関」で、いくら以下だと「低い相関」だというようなことを一般的に述べることはできない。ただ、次に述べる検定(相関係数の検定)を用いて、「有意でない相関係数」は低い相関だとみなしていいだろう。

(「社会調査法入門(2004年、盛山和夫著)」より)

社会事象は、コンテキスト依存性^{*1}が強い。変数と変数の間に関連性がある方がふつうなのである(盛山・近藤・岩永, 1992:99)。そのため、得られた関数が「意味ある」ためには、一定水準で有意というだけでなく、その絶対的な大きさが必要になる。…(中略)…その大きさの閾値^{*2}が学問分野によって違う。経済学では相関係数0.9以上(決定係数0.8以上)が要求されることが多いが、社会学では0.4~0.5ぐらいが目安になる。

(「実践としての統計学(2000年、佐伯胖、松原望編)」より)

- *1 コンテキスト依存性: 文脈依存性。「社会事象は、さまざまな他の事象や社会的な背景等に影響されやすい」という趣旨であると思われる。
- *2 閾値: ある系に注目する反応をおこさせるとき必要な作用の大きさ・強度の最小値(広辞苑第六版)。ここでは、その値を境にして意味が変わる値のこと。

(3) 回帰係数、決定係数

一方の変数の値から他方の変数の値を予測したり、一方の変数の変化が他方の変数に与える影響を調べたりするなど、2つの変数の役割をはっきり区別して、その関係を調べる方法の一つが回帰分析。2変数(x 、 y)のデータを対象とする単回帰分析では、観測データから「 $y = ax + b$ 」で表される回帰直線を推定し、予測等を行います。この a が回帰係数(=直線の傾き)です。

このとき、 y が x によってどれくらい説明されるか(x が y をどの程度決定しているか)を表す指標が決定係数です。決定係数が、1に近いほど観測値(計算に用いた実際の数値)と予測値(式に当てはめ計算した値)が近いことを意味します。回帰係数は関係そのものを表し、決定係数は関係の強さを表すと言えます。

