

[演題3]

統計分析から見た人の個性についての考察

高見正利

社会リハビリテーション学科

統計分析は客観的なデータ処理法であり、数学の手法を用いてデータを解析して数値上の性質や規則性などの情報を提供する。経験的に得られたばらつきのあるデータから、偶然の変化（誤差）なのか、何か意味ある変化なのか、を確率で判断する。なかなか起きにくい確率とする5%や1%

の領域に入ると、これは何かあるぞ、普段と違うと警戒するわけである。代表的なものとして、相関分析、t検定、分散分析、カイ二乗分析などがある。判断のもとになる確率の分布が、処理の種類により異なるので、それを使い分ける必要はあるが、これらはみんな親戚関係にある。



図1A 散布図A
(相関が強い)

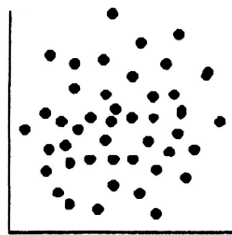


図1B 散布図B
(相関は無い=x、y軸は独立)

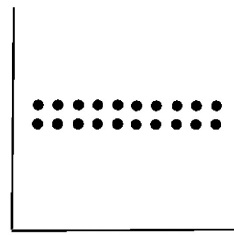


図1C 散布図C

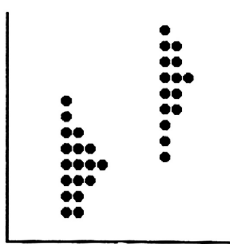


図2A t検定A
(有意差あり)

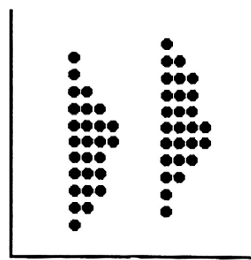


図2B t検定B
(有意差なし)

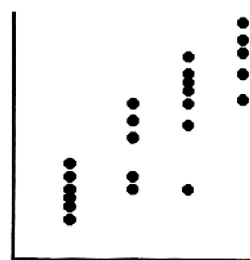


図3A 分散分析A
(有意差あり)

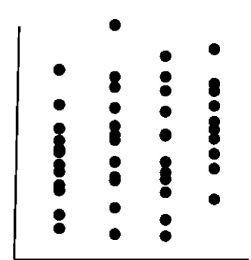


図3B 分散分析B
(有意差なし)

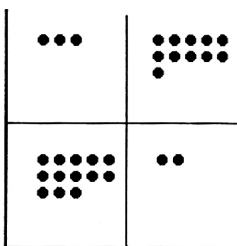


図4A カイ2乗分析
(有意差あり)

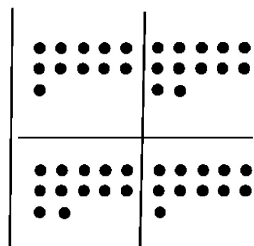


図4B カイ2乗分析
(有意差なし)

1. 相関分析

まず、データが得られると、散布図や階級に分けてヒストグラム(度数分布図)にまとめられる。散布図を例に示すが(図1A、図1B、図1C)、これは横軸(x)および縦軸(y)が調査項目(変数)に対応して値が配置される。身長、体重などの間隔・比率尺度、(順序尺度)のデータが該当する。その際に、よく使われるのが相関分析である。これは、2変数間の関係を分析する方法で類似性の度合いを相関係数という数値で示す。変数間の関係を調べることは頻繁にあり、2つのデータの関連性が強ければ相関係数は1に近づき、関連性が低ければ0に近づく。この相関を判断する相関係数検定表はt分布表をベースに作られている。

2. t検定

横軸の項目について2つのレベルに分けると(背の高い群と低い群などと)、t検定になり、縦軸は間隔・比率尺度のまま横軸は名義尺度となる。2組のデータ群に違いがあると言っているのかを判断する。すなわち2群の平均値の差が個々のデータのばらつきの主要な原因になっているのかどうかをt分布に照らしてみ、その差がある値なのかをチェックするのである。図2Aは有意差あり、図2Bは有意差なし、の状況を表している。また、t分布は正規分布の近似である。

3. 分散分析

横軸の2つのレベルだけでなく3つ以上にすると分散分析と呼ばれるものになる。単なる違いだけでなく、少し詳しい分析ができる。何が影響しているかがわかるようになり要因間の相乗効果も判断することができる。データのばらつき(誤差)

が、偶然伴ってでてくる意味のないばらつき(誤差変動)なのか、要因によって変化した意味のあるばらつき(要因による変動)なのかを判定する。ばらつきの大きさの比をとり、F分布と呼ばれる確率分布に照らして、要因の効果を判断する。要因によるばらつきが誤差に比べて大きければ、要因による影響があるとする。図3Aは有意差あり、図3Bは有意差なし、の状況を表している。また、F分布は下記のカイ2乗分布から得られる。

4. カイ2乗検定

今度は、縦軸も単純化して2つのレベルに分割すると、カイ2乗検定あるいは独立性の検定、適合度の検定と呼ばれるものになる。データが予測される割合通りになっているか、あるいは偏りがあるのかどうか、2つの変数についての集計表(分割表)にある偏りがそれぞれに関連しているのか、を判定する。図4Aは有意差あり、図4Bは有意差なし、の状況である。参照されるのがカイ2乗分布であり、正規分布がベースになっている。つまり、ここで挙げた確率分布はすべて正規分布とつながっている。

5. 考察

さて、どれをとってもデータのばらつきや偏りという言葉(分散とも呼ばれる)が現れた。データのばらつきがなければ何も言えない、情報が無いということになる。データが1か所に(同じ値に)固まっていたり、均等に分布していたりすると抽出する情報がないことになる。データにはばらついていて欲しい。情報を得たい者にとっては、むしろ癖のある分布の方がありがたい。ばらつきがあって初めて関心もたれ、そのばらつきがどのような要因で起きるのか、それを確率で判断していく。そして、有用な知見(情報)が見つかる

のである。

こう考えていくと、人間社会も同じではないか。いろいろな特徴や意見を持つ人がいて、お互いの足りないところを補いあい、またそれだけでなく、普段は目立たなくとも、イザという時には俄然活躍することになる。そのような具合で人間社会が成り立っていく。環境が違う、性格が違う、価値観が違う、目指すものが違う、行動が違う、好きなことが違うなど様々な人がいた方がいいのではないか。自分では経験できなかったことや考えに至らなかったことなどに触れて、刺激を得ることができ、世界観が広がる。さまざまな個性は豊かな情報をもっているのである。自分とまったく同じ人間ばかりだとどう感じるだろうか。少なくとも面白味はない。他人と同じでないことを喜ぶべきであり、違うところで自分の良さを発揮できる。統計学に触れて、はからずも個性尊重という理念を深めることができた。