

連載「歯科心身医学研究に役立つ統計学講座」

第3回 関連を検討する際のポイント

松岡 紘史

Statistics lecture for the study of psychosomatic dentistry 3rd point for analysis clarifying correlation

Hirofumi Matsuoka

Abstract: There are a number of important points to enhance the reliability for conducting statistical analysis and interpreting the results. In this article, these vital points for statistical analysis clarifying correlations were explained based on the (1) significant of analysis and interpretation of correlation coefficient, (2) outlier and scatter diagrams, and (3) partial correlation coefficient.

key words: correlation coefficient, scatter diagram, partial correlation coefficient

キーワード：相関分析，偏相関分析，散布図

連載2回目で扱った差を検討する統計解析¹⁾とともに、非常に頻繁に扱われる仮説は2変数の間に関連があるかというものである。本論文では、関連を検討する解析において注意すべき点について述べる。

1. 相関係数の有意性と大きさの解釈

変数間の関連を検討する際に、もっとも広く用いられるのは、ピアソンの積率相関係数やスピアマンの順位相関係数である。これらの分析を実施すると、変数間の関係が相関係数とp値として表される。この際、差を検討する際の注意点として述べた検定の有意性と効果サイズ¹⁾が、相関分析でも重要となる。相関係数で検討される有意性とは、一般的に無相関検定と呼ばれ、相関が見られない（相関係数がゼロである）ことを前提とした場合に、両者にみられた関連が偶然生じる可能性を表しており、p値が0.05未満の場合は、変数間に何らかの関連が認められることを意味している。また、相関分析で算出される相関係数では、関連の強さを-1～1の数値で表す。相関係数の大きさは、0.00～±0.20は「ほとんど相関はない」、±0.20～±0.40は「弱い相関がある」、±0.40～±0.70は「比較的強い相関がある」、±0.70～±1.00は「強い相関がある」として解釈される²⁾。このため、相関係数が有意な値であってもその値が0.20未満であれば、変数

間にはほとんど相関がみられないということになる。ただ、こうした相関係数の大きさの解釈はあくまで目安であり、相関係数は2乗した場合に、一方の変数が他方の変数の変動を説明する割合を意味しているため、実質的にどの程度の説明割合があるかを基準に、変数間の関係に意味があるかを解釈する必要がある。

2. 相関係数算出の際に注意が必要な点

相関係数を算出する際の注意点の1つは、外れ値の存在である。さきほどあげたピアソンの積率相関係数とスピアマンの順位相関係数のうち、外れ値が存在した場合の影響が大きいのはピアソンの積率相関係数である。例えば、図1に示したような分布の場合、外れ値を含まない相関係数は0.70程度であるが、外れ値を含んだ場合は、相関係数が大きく変化し、その変化はピアソンの積率相関係数の方が大きい（ピアソンの積率相関係数：0.33、スピアマンの順位相関係数：0.53）。また、この外れ値の影響はサンプルサイズが小さい場合に大きくなる。こうしたことから、外れ値によって分析結果が影響を受ける可能性がある場合は、分析方法としてスピアマンの積率相関係数を選択することが望ましい³⁾。なお、外れ値を検出する方法としては、スミルノフ・グラブス検定や箱ひげ図を利用する方法がある。

北海道医療大学歯学部口腔構造・機能発育学系保健衛生学分野

Department of Oral Growth & Development, Division of

Disease Control & Molecular Epidemiology, School of Dentistry, Health Sciences University of Hokkaido

(受付日：2017年7月25日)