

対数変換が主成分分析に及ぼす影響

お茶の水女大家政 山路博子

目的 集団の特性を把握するためには様々な統計的手法が施されるが、これらはデータの分布状態を十分に考慮した上で用いられなければならない。主成分分析を行なう場合においても、その分布が非対称のときには、対数変換を行なった方が解析結果が明確になることがある。そこで今回は、資料の分布型の検定方法、および対数変換が主成分分析に及ぼす影響について検討を行なった。

資料・方法 資料は昭和46・47年の体格調査により得られた成人女子（25-69歳）、4123名の身体計測値であり、研究項目は高径5項目、長径6項目、周径8項目、幅径1項目、皮脂厚2項目と体重の23項目である。解析は全年齢を一括して行なった。なお、分布型は尖度、歪度により検定した。

結果 (1) 級間のとり方で尖度、歪度が変化してしまいますが、級間の幅を1にして分布型の検定を行なうことと、このような誤差を防ぐことができた。(2) 正規分布を示さない項目は対数変換を行なうことと正規分布に近づくが、逆に正規分布を示していた項目は正規分布を示さなくなる傾向が計られた。(3) 正規分布を示さない項目を対数変換することで基準化した後に主成分分析を行なった結果と、対数変換を行なわない場合の結果を比較すると、固有値、累積寄与率、因子負荷量とともに、両者の間にほとんど差が計られなかった。すなわち、成人女子を対象に主成分分析を行なう場合、対数変換による基準化は必要ないという結論を得た。