

卒業研究概要

成績：

提出年月日 2009年 1月 29日

卒業研究課題 主成分分析と因子分析による競馬の勝因の研究

学生番号 A04-133

氏名 邊見 広大

概要 (1000字程度)

指導教員 真貝 寿明

印

主成分分析と因子分析を用いて、毎年12月に中山競馬場で行なわれる有馬記念のレース結果を予想した。主成分分析とは、勝因と考えられるいくつかの要素を線形結合し、相関行列の固有ベクトルを用いる分析である。因子分析とは、任意の主要素を求める際に、各要素の残余項を最小にするように求める分析である。主成分分析と因子分析の結果から、競馬で予想の時に用いられるデータのどれが実際に影響して良い成績を残したのかを調べ、2008年の有馬記念の出走馬に適用した。

データは、過去4年分の有馬記念の出走馬の成績と、過去4年分の中山競馬場2500mのレースで勝った馬の成績を標準化して用い、成績データは、右回り、左回り、競馬場、距離、GI、重賞競争の成績(1着、2着、3着、着外)の結果から3要素を決め、主成分得点、因子得点を計算するプログラムをC言語で作成した。要素を変えながら図示し、最も相関があると考えられる3要素を探した。

図1、2は過去の有馬記念から、右回りと競馬場を組み合わせ、距離、GIと重賞組み合わせた3要素をとり主成分得点と因子得点の図示し、2008年の有馬記念を予想に用いたものである。

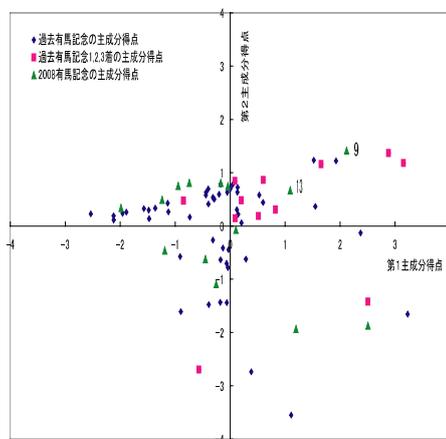


図1. 主成分分析の散布図

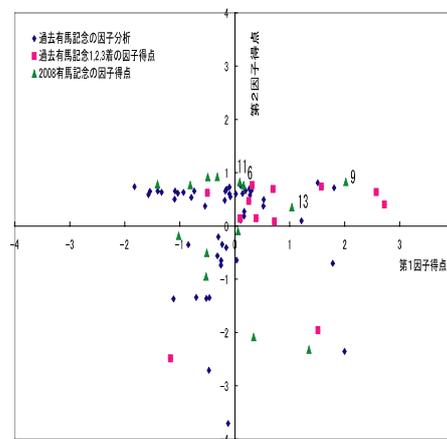


図2. 因子分析の散布図

主成分分析と因子分析は同様に分布するが、主成分分析はデータが楕円状に分布し、因子分析はデータが円状に分布する事がわかった。どちらも分析固有の性質が明確に出た。図1の第1主成分得点は3つの要素が高い馬がプラスを表し、第2主成分得点はマイナスなら距離の要素が高く、プラスなら残りの要素が高いことを表した。図2の第1因子得点は、距離以外の要素で説明でき、第2因子は距離の要素だけで説明できた。

主成分、因子分析ともに右上のほうに集まるように相関が見えるため、2008年の有馬記念の予想は馬番の順に6番、9番、11番、13番で、特に主成分、因子ともに右上に来ている9番と13番が勝つ確率が高いという結果になった。実際のレースの結果は1着から13番、14番、6番の順であった。研究結果から、2頭入選したことになった。