Analysis of the Profile of Golfers using Multivariate Methods

Kuldeep Kumar

School of Information Technology

Bond University

Gold Coast, AUSTRALIA

kkumar@bond.edu.au

J.H. Park

School of Information Technology

Bond University

Gold Coast, AUSTRALIA

jhpark@bond.edu.au

1. Introduction

With the advancement of World Wide Web and Internet technology data on golf is available in abundance. Statistical techniques like probability theory and correlation and regression analysis (Jones 1992), and time series analysis (McCullough 1992), are used to study the statistical profiles of the golfers. Multivariate technique deals with several variables simultaneously and has proved to be successful in analysing sports data (See Ganesalingam and Kumar 1994). However, very little work has been done regarding the application of the technique in the analysis of golf data.

The multivariate techniques employed in this study are multiple regression analysis, principle component analysis and factor analysis. The aim is to identify the factors, which are significant in explaining the performance of the golfers. In this study we have collected the data from on several variables, which are helpful in explaining the performance of the golfers. This data has been collected from the statistical summary of the top 30 players in the 1999 U.S. Open Championship.

2. Methodology

First of all, we have used the technique of multiple regression analysis including stepwise regression analysis to see which variables are significant in explaining the performance of individual golfers. When a large number of measurements are available, it is natural to enquire whether they could be replaced by less number of measurement or of their function, without loss of much information for convenience in the analysis and in the interpretation of the data. Principle component, which are linear function of the measurement, are suggested for this purpose. Another technique, which is commonly used in multivariate analysis, is factor analysis. Through factor analytic techniques, the number of variables for further research can be minimised while also maximising the amount of information in the analysis. Riccio (1994) has used factor analysis to determine which part of the play most significantly affect golfer's decline in competitiveness due to age. Mathematical details of these multivariate techniques can be found in Flury (1997).

3. Analysis and Results.

From the multivariate analysis we conclude that the average driver distance of 4 rounds of golf, the average number of putts made in 4 rounds of golf, the number of greens hits in 4 round golf out of 72 holes and the average number of putts made per hole in 4 rounds of golf are the significant variables in explaining the performance of a golfer which we have measured by total score in four round of golf. We have also done Principal Component and factor analysis, which could be helpful for an amateur golfer to compare himself with the professional golfer.

One of the advantages of the techniques used in this paper is to help those golf players who want to improve their skill's level by benchmarking professional players on PGA tour. It also would be expected to motivate more players into the game. According to the research conducted by Beditz (1994) the time and money became the considerable major barriers in golf participation. Implementing the techniques used in this paper to teaching areas would expect to reduce those major variables by Beditz and increase the motivation of the players and game's participation rates.

RFERENCES

1. Beditz, J.F. (1994). " The development and growth of the U.S. golf market" in Cochran, A.J., and M.R. Farrally, eds. Science and Golf: Proceedings of the World Scientific Congress of Golf. London: E & FN Spon.

2. Flury, B (1997) " A first course in Multivariate Statistics", Springer Text in Statistics

3. Ganesalingam, S. and Kumar,K. (1994). " A Statistical look at cricket data" in Neville De Mestro eds, Mathematics and Computers in Sport Vol. 2,pp 89-104.

4. Jones, R. E. (1992). "A Correlation analysis of the Professional Golf Association (USA) Statistical rankings for 1988" in Cochran, A.J. Science and Golf: Proceedings of the World Scientific Congress of Golf. London: E & FN Spon.

5. McCullough, J. M. (1992). "The application of time series decomposition techniques to the analysis of golf performance" in Cochran, A.J. Science and Golf: Proceedings of the World Scientific Congress of Golf. London: E & FN Spon.

6. Riccio, L. J. (1992). " Statistical analysis of the average golfer" in Cochran, A.J. Science and Golf: Proceedings of the World Scientific Congress of Golf. London: E & FN Spon.

RESUME

Dr Kuldeep Kumar is Associate Professor of Statistics at Bond University. He has presented papers in the Conference on Mathematics and Computers in Sport. In 1998 he was one of the Director of the above conference. John H Park is a professional golfer and one of the research students.