

## Curving the golf ball: Predicting the renewal of golf club memberships

Abstract ID: EASM-2015-251/R1 - (777)

All authors:

Katharine Hoskyn (corresp), Michael Naylor, Geoff Dickson

Date submitted: 2015-03-25

Date accepted: 2015-04-12

Type: Scientific

Keywords: Membership renewal

Logistic regression

ROC curve

Category: 1: Critical Sport Management issues

Synopsis:

The aim of this research is to investigate the value of membership data in predicting whether or not players renew their annual subscription to member-owned golf clubs in New Zealand. Identifying the characteristics of at-risk golfers will assist in the development of membership retention strategies and initiatives. This study uses logistic regression and a ROC curve in analysis of membership data for New Zealand golf clubs over a four year period. Players who continued their membership were compared with those who discontinued their membership at some time between 2009 and 2012 (n=2226). Younger players, with no handicap and no linked friends are more at risk discontinuing membership.

Abstract:

AIM

The aim of this research is to investigate the value of membership data in predicting whether or not players renew their annual subscription to member-owned golf clubs in New Zealand. Most New Zealand golf clubs have experienced member decline for the past decade. Identifying the characteristics of at-risk golfers will assist in the development of membership retention strategies and initiatives.

LITERATURE REVIEW

Predictive risk modelling has been used by the medical profession to identify patients at risk of medical conditions. One approach is to use the result of logistic regression analysis in conjunction with a receiver operating characteristic (ROC) curve. Logistic regression is appropriate when the dependent variable has two mutually exclusive categories, in this case continuing or discontinuing membership. The technique uses independent variables to predict the probability that a person will belong to one of the two categories (Tabachnick & Fidell, 2013).

Logistic regression has been used in sport to identify variables predicting membership of sport groups. Downward and Rasciute (2010) summarise nine studies in which logistic regression was used to identify predictors of sport participation. Four of these studies used logistic regression to predict participation in sport activity groups, such as club-organised sport or some infrastructural group such as a gym. None of these studies specifically investigate club membership. More recently McDonald, Karg and Leckie (2014) used logistic regression in their investigation of season ticket holders at risk of not renewing their season ticket.

One method for assessing the strength of a logistic regression model is to consider its success in correctly predicting the outcome for cases when the actual outcome is known. Logistic regression produces a classification table of incorrect and correct classifications. In biomedical application, specificity and sensitivity refer to the correct classification of cases (Tabachnick & Fidell, 2013). Sensitivity is the correct classification of discontinuing members and specificity the correct classification of continuing members. The ROC curve takes a visual approach to sensitivity and specificity. Agresti (2007) maintains that it is more informative than a classification table. In addition, it conveys predictive power at a glance.

#### METHODOLOGY, RESEARCH DESIGN AND DATA ANALYSIS

This study uses logistic regression and a ROC curve in analysis of membership data for New Zealand golf clubs over a four year period. The dataset includes age, gender, handicap (an indication of ability); total frequency of playing; frequency of playing at home club; and the number of friends to whom a player is linked in the membership system. The data was analysed in two ways. First, we analysed all members who discontinued their membership at some time during the four-year period. Second, we conducted a longitudinal analysis of members who discontinued in the final year of analysis.

A key limitation of the study is that analysis is restricted to variables in the membership data set. The format of some variables also placed limitations on the analysis that could be undertaken. In addition, multicollinearity (high correlation) between several variables limited the variables suitable for the logistic regression model.

#### RESULTS, DISCUSSION AND CONCLUSIONS

Players who continued their membership were compared with those who discontinued their membership at some time between 2009 and 2012 (n=2226 members from 308 of the 393 golf clubs in New Zealand). The final logistic regression model included three variables (age, the number of friends linked to the player in the membership database and whether or not the player had a handicap). Younger players, with no handicap and no linked friends are more at risk discontinuing membership. The model containing these predictors was statistically significant. Chi-square (3, n= 2226) = 1180.324,  $p < .001$ . This indicates that the model distinguishes between members who continued and those who discontinued membership. The model explained between 41.3% (Cox and Snell R squared) and 55.1% (Nagelkerke R squared) of the variance in membership behaviour, and correctly classified 83.4% of total cases. The correct prediction rate was high for both membership groups: 79.3% of discontinued memberships and 87.2% of continued memberships were

correctly predicted. All of the variables made a statistically significant contribution to the model. Further analysis using logistic regression and descriptive statistics with three years of longitudinal data shows a similar pattern. The ROC curve visually demonstrates the strength of this model by plotting correct classification on a graph.

Of equal importance, the analysis demonstrates variables that are not predictors of discontinuing membership, such as gender and level of handicap. The relationship between frequency of playing and membership continuation is complex, as shown by the longitudinal analysis over three years.

This study empirically estimates indicators of discontinuing membership from a membership database. Results can potentially be used by golf club managers to identify at-risk members and shape subsequent interventions.

References:

Agresti, A. (2007). *An introduction to categorical data analysis* (2nd. ed). New Jersey: Wiley

Downward, P., & Rasciute, S. (2010). The relative demands for sports and leisure in England. *European Sport Management Quarterly*, 10:2, 189-214. doi.org/10.1080/16184740903552037

McDonald, H., Karg, A.J., & Leckie, C. (2014). Predicting which season ticket holders will renew and which will not. *European Sport Management Quarterly*, 14:5, 503-520. doi.org/10.1080/16184742.2014.944199

Tabachnick, B.G., & Fidell, L.S. (2013). *Using multi-variate statistics* (6th ed.). Boston: Pearson Education.