International Journal of Golf Science, 2013, 2, 152-161
http://dx.doi.org/10.1123/ijgs.2013-0014


INTERNATIONAL JOURNAL OF
GOLF SCIENCE
Official Journal of WSCG
www.IJGS-Journal.com
ARTICLE

# Applying the Hardy Distribution to the Hole Scores of the 2012 British Open Championship

**A.H.G.S. van der Ven**
Radboud University

In an article titled "A Mathematical Theorem about Golf", Godfrey Harold Hardy (Hardy, 1945) introduced a simple model of golfing. Hardy models golf as a sequence of independent shots, each of which can be ordinary (or normal), good (or excellent) or bad. In a more recent paper titled "The Hardy Distribution for Golf Hole Scores" van der Ven (van der Ven, 2012) derived the probability distributions of scores for par three, four and five holes according to Hardy's basic assumptions. In this paper the results are reported of a goodness-of-fit test of Hardy's theorem using the hole-by-hole scores of the 2012 British Open Championship. Pearson's chi-square goodness of fit test was used to determine whether the observed sample frequencies of the hole scores differed significantly from the expected frequencies according to Hardy's model. The fit between observed and expected frequencies was generally very satisfactory.

***Keywords:*** probability theory, Hardy distribution, golf hole scores

The frequency distribution of hole scores can be obtained for each hole after each round (or each day) of a professional golf tournament. For example, after the first round of the British Open of 2012 the frequency distribution of the scores for the first hole, a par three, was as follows:

| **Score** | 2 | 3 | 4 | 5 |
|---|---|---|---|---|
| **Frequency** | 21 | 117 | 17 | 1 |

The question arises whether it is possible to develop a model to explain the particular shape of this distribution. In an article titled "A Mathematical Theorem about Golf" G.H. Hardy (Hardy, 1945) introduced a simple model, which was based on the following assumptions:

i For one hole a golfer has probability $p$ of gaining a stroke with a single shot and
ii probability $q$ that his/her shot costs him a stroke.

---

van der Ven is with the Department of Learning and Development, School of Pedagogical and Educational Science, Radboud University, Nijmegen, The Netherlands.

According to Cohen (Cohen, 2002):

> Such strokes will be described as good (*G*) or bad (*B*), respectively, leaving probability 1 - *p* - *q* of an ordinary (*O*) stroke. Then, for example, on a par four hole, successive strokes *OGO* will result in a birdie (a score which is one stroke less than par) and *BBGOO* in a bogey (a score which is one stroke more than par).

This distinction into three different kinds of strokes is completely in agreement with a distinction made by Broadie in 2008 (Broadie, 2008, p. 3), who distinguishes between an *awful* shot, a *great* shot and a *typical* shot:

> An *awful* shot is any shot with a shot value[1] less than –0.8; a *great shot* has a shot value greater than +0.8. A *typical* shot is any shot that is neither awful nor great. Examples of awful shots include any shot hit out of bounds, most shots hit into water, hitting a 60-yd shot into a greenside bunker, hitting a tree on a drive so the ball only travels 70 yds, and missing a 2-ft putt. Examples of great shots include hole outs from off the green, hitting a 90-yd shot to 2 ft from the hole and hitting a 210-yd shot to within 10 ft of the hole. (p. 3).

More recently, van der Ven (van der Ven, 2012) published a paper in which he derived the probability distribution $P(T_m=n)$ for the number of strokes *T* (*T* from *Total*) a player may take on a hole of par *m*. Separate formulas were derived for the probability distribution of

a par three hole: $P(T_3=n)$, n = 2, 3, 4, …

a par four hole: $P(T_4=n)$, n = 2, 3, 4, …

a par five hole: $P(T_5=n)$, n = 3, 4, 5, …

For example, a birdie on a par three hole can be obtained with the following sequences of strokes

OG with probability $(1-p-q)\,p$

GO with probability $p\ (1-p-q)$

GG with probability $p^2$

Therefore, for $T_3$=2 one obtains:

$$P\ (T_3=2) = 2\,p\ \ (1-p-q)+\mathrm{p}^2$$

Note, that according to Hardy's model for $m = 3$ a score of 1 (a so-called 'hole in one') is impossible. Therefore, in Hardy's model a 'hole in one' is treated as a birdie. Similarly, for $m = 5$ a score of 2 (a so-called 'albatros') is impossible. Therefore, in the case of a par five, an albatross is treated as an eagle (a score of two under par).

In this paper the hole-by-hole scores of the 2012 British Open Championship were used to empirically test Hardy's model by comparing the observed frequencies of hole scores with the expected frequencies under Hardy's assumptions. Hardy's model however, applies only for the case of a single player. If the scores are obtained from different players then the performance of a model fit would only make sense if all players have equal skill. This problem will be dealt with in the sections: Equivalence of Players and Testing Equivalency. The results of the actual goodness-of-tests are reported in the section: Goodness of Fit Tests. The tests were only applied to the results of round 1 and round 2. In comparison with the number of players participating in round 1 (day one) and round 2 (day two) the number of players participating in round 3 (day three) and round 4 (day four) was considerably reduced. Finally, some of the results are discussed in more detail.

## Equivalence of Players

To ascertain the validity of a certain probability distribution it is common practice to compare the observed frequency distribution of scores with the expected frequency distribution using a goodness-of-fit test such as, for example, the Kolmogorov-Smirnov Z test or Pearson's chi-square test. To do so, however, one must have at his/her disposal many scores. It is very difficult, however, to find a player prepared to play a hole many times to obtain sufficient scores for such an analysis. Moreover, due to practice, the player might become more familiar with the hole in the course of playing. This would mean, that subsequent hole scores would undergo some learning effect and could, therefore, not be considered as a collection of pure replications. The latter criteria are a necessary condition for the application of a goodness-of-fit-test, which is based on a comparison of the observed and expected frequency distribution. A possible way to circumvent this problem would be to use different players, each player playing the hole one time. However, this only would make sense if there is sufficient evidence that the players (competitors) would all perform at the same skill level.

In analogy with the concept of *replicate measurements* in test psychology, (see Lord and Novick, 1968, chapter 1.12, page 46) one may consider the first two rounds of a tournament as two replicate tests. This would also be the case for the second two rounds. However, in the case of these rounds the number of players is considerably reduced, which make these rounds less suitable for a goodness-of-fit test. According to the *classical test theory model* (see Lord and Novick, 1968, part 2, page 55) the correlation between replicate tests is a measure for the reliability of the test. The reliability of the observed test (or round) score $X$, which is denoted as $\varrho_{XX'}$ is defined as the ratio of true score variance $\sigma_T^2$ to the observed score variance $\sigma_X^2$:

$$\varrho_{XX'}^2 = \frac{\sigma_T^2}{\sigma_X^2} = \frac{\sigma_T^2}{\sigma_T^2 + \sigma_E^2}$$

where $X$ and $X$' are replicate tests (round 1 and round 2). The observed score $X$ equals the true score $T$ plus some error $E$ *i.e.,* $X=T+E$. Therefore, if the correlation between round 1 and round 2 is zero (or not significantly different from zero), one may validly conclude that the systematic differences between players may be neglected, and the round scores may be considered as being produced by an imaginary 'single' player. In a study of Clark (Clark, 2004), which involved, among others, professional golfers in 1999 on the Professional Golfers' Association Tour (PGA Tour), low correlations were found between round 1 and 2. When all players, who played in the first two rounds of tournaments with four rounds, were included, the average correlation between players' rounds 1 and 2 scores was equal to 0.22 ($N$ = 5,481). According to Clark (2004):

> With Ns greater than 2,500, any correlation of .03 would be statistically significant. However, since virtually any study can be made to show significant results if one uses enough subjects, and statistical significance says nothing about the magnitude or importance of the relationship between two variables, we followed Cohen's general rule (1977) for deciding whether a correlation is weak or strong. The qualitative descriptors of small ($r$ = .10 - .29), moderate ($r$ = .30 - .49), and large ($r$ = .50 +) were applied here. (p. 676).

Therefore, even if all participants of a professional golf tournament are included in the analysis, there still exists strong evidence for low correlations between round 1 and round 2. What has been said here for round scores, naturally also holds for hole scores.

## Testing Equivalency

The goodness-of-fit tests to be reported below were all based on the hole-by-hole scores of rounds 1 and 2 of the 2012 British Open Championship. These data were also used to test the assumption of equivalency of players. For any goodness-of-fit test the sample size has to be sufficiently large. Therefore, the decision was made only to use the results of round 1 and 2 for a goodness-of-fit test and to perform the test for round 1 and round 2 separately as a double check. In the British Open Championship of 2012 the correlation between the total score of round 1 and 2 was significant at the 0.01 level ($r$ = .217, 2-tailed $p$ = .007, $N$ = 156). One of the players, Mardan Mamat, was disqualified after round 2 when it was discovered that he had signed an incorrect scorecard. He was disqualified for signing for a lower score at the sixth hole, which he penciled down as a birdie three instead of a four. He also signed for a five on the eighth hole when, in fact, he had shot a par four. For the statistical analyses made in this paper, however, he was included in the data set with the use of his correct scores. This explains why $N$ = 156. Although the correlation between the total score of round 1 and 2 was very low, it was never-the-less significant. Therefore, one might argue that these data cannot be used for a goodness-of-fit test, because there is at least some evidence against the assumption of equivalence of the players.

A possible way to circumvent the problem of a significant correlation between round 1 and round 2 scores would be to take into consideration that the goodness of fit tests to be used are performed on hole scores instead of on round scores. Therefore, instead of inquiring about the relationship among round scores, the relationship among hole scores was focused upon. If there are no systematic ('true') differences between players, then, for each hole, the correlation between the hole score in round 1 and the hole score in round 2 should not deviate significantly from zero. These correlations are given in Table 1 (see column 6). The two-tailed probabilities are given in the next column. Although the distributions of the hole scores are not normal, the $t$ test used is both powerful and sufficiently robust against nonnormality (see Edgell and Noon, 1984). The observed correlation deviates significantly from zero only in the case of the first hole, at the 5%-level. The correlation between the hole scores, however, is still very low ($r$ = .170) also according to Cohen's general rule (Cohen, 1977). Because for the remaining holes, (i.e., holes 2–18) the correlations are not significant, one may soundly conclude that, at

**Table 1 Paired Sample *t* Test for the Difference Between the Means per Hole of Round 1 and Round 2 (*N* = 156, *pr*: Two-Tailed Probability for $H_0$ *r* = 0 and *pt*: Two-Tailed Probability for $H_0$ *t* = 0).**

| | Round 1 | | Round 2 | | Correlation | | Paired *t* test | |
|---|---|---|---|---|---|---|---|---|
| *Hole* | *m* | *s* | *m* | *s* | *r* | $p_r$ | *t* | $p_t$ |
| 01 | 2.99 | 0.520 | 2.95 | 0.577 | 0.170 | 0.034 | 0.678 | 0.499 |
| 02 | 4.06 | 0.608 | 4.16 | 0.585 | -0.011 | 0.892 | -1.416 | 0.159 |
| 03 | 4.46 | 0.782 | 4.56 | 0.772 | -0.033 | 0.686 | -1.219 | 0.225 |
| 04 | 4.06 | 0.593 | 4.01 | 0.709 | -0.048 | 0.553 | 0.592 | 0.554 |
| 05 | 3.00 | 0.556 | 3.14 | 0.606 | 0.019 | 0.813 | -2.162 | 0.032 |
| 06 | 4.54 | 0.790 | 4.41 | 0.641 | 0.071 | 0.382 | 1.631 | 0.105 |
| 07 | 4.76 | 0.774 | 5.11 | 0.783 | -0.009 | 0.910 | -3.981 | 0.000 |
| 08 | 4.22 | 0.712 | 4.26 | 0.818 | 0.103 | 0.202 | -0.467 | 0.641 |
| 09 | 2.91 | 0.594 | 2.90 | 0.752 | 0.038 | 0.635 | 0.085 | 0.932 |
| 10 | 4.12 | 0.672 | 4.10 | 0.654 | -0.130 | 0.106 | 0.161 | 0.873 |
| 11 | 4.96 | 0.726 | 5.03 | 0.652 | -0.039 | 0.631 | -0.805 | 0.422 |
| 12 | 3.10 | 0.465 | 3.10 | 0.634 | -0.055 | 0.491 | -0.099 | 0.921 |
| 13 | 3.86 | 0.647 | 3.84 | 0.606 | 0.008 | 0.923 | 0.272 | 0.786 |
| 14 | 4.13 | 0.663 | 4.05 | 0.620 | 0.077 | 0.338 | 1.193 | 0.235 |
| 15 | 4.32 | 0.727 | 4.39 | 0.758 | 0.122 | 0.128 | -0.895 | 0.372 |
| 16 | 3.83 | 0.708 | 3.86 | 0.550 | -0.028 | 0.732 | -0.353 | 0.725 |
| 17 | 4.24 | 0.616 | 4.09 | 0.722 | 0.038 | 0.641 | 2.064 | 0.041 |
| 18 | 4.03 | 0.704 | 4.04 | 0.756 | -0.100 | 0.216 | -0.148 | 0.883 |

the level of hole scores, systematic differences between players may be neglected. In addition to the correlations between the hole scores obtained in round 1 and round 2 one may also take into consideration the differences between the means of these hole scores. The results of the paired sample *t* tests are displayed in the last two columns of Table 1. In the majority of the holes the differences were not significant. A significant result was only found for hole 5 and 17 at the 5%-level and for hole 7 at the 1%-level. In major golf tournaments the difficulty of some holes is increased progressively from rounds 1 through 4 by tournament organizers through altering components of holes, most notably, the pin position on the green. Further, British Open tournaments which are typically played on links courses are prone to considerable environmental extremes (specifically relating to wind).

## Goodness of Fit Tests

Maximum likelihood estimation (*MLE*) for *p* and *q* was performed by maximizing the log-likelihood *logL*:

$$logL = \sum f_n \ln\left(P\left(T_m = n\right)\right).$$

where $f_n$ refers to the observed frequency belonging to a score of $n$. Each of the partial derivatives of $logL$ with respect to $p$ and to $q$ was set equal to zero and the resulting two equations were solved for $p$ and $q$. Solving was done numerically using the function *fsolve*[2] from Maple 12[3]. The function *fsolve* requires the implementation of initial values. The moment estimates for $p$ and $q$ were used as initial values. If the moment estimates were negative the initial values were set equal to a value close to zero (0.0001). The maximum likelihood estimates of $p$ and $q$ are denoted as $\hat{p}$ and $\hat{q}$. The values of $\hat{p}$ and $\hat{q}$ are given in Table 2.

The $\chi 2$ goodness-of-fit test was used as a model test. The $\chi 2$ goodness-of-fit test is based on a comparison of the observed frequencies with the expected frequencies. The expected frequencies were computed with the estimated parameters $\hat{p}$ and $\hat{q}$. The actual test was performed as follows. For each hole the observed frequencies for a birdie (one under par) or lower were taken together. The same was done for the corresponding expected frequencies. Likewise the observed frequencies for a double bogey (two above par) or a higher score were taken together, which was also done with the corresponding expected frequencies. As a result, for all holes, the

**Table 2 Estimated Values of $p$ and $q$ and $\chi^2$ Test Statistics ($N$ = 156, Length in Yards, $p_{\chi 2}$: Right Tail Probability for $\chi^2$).**

| | Properties | | Round 1 | | | | Round 2 | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Hole | Length | Par | $\hat{p}$ | $\hat{q}$ | $\chi^2$ | $P\,\chi^2$ | $\hat{p}$ | $\hat{q}$ | $\chi^2$ | $P\,\chi^2$ |
| 1 | 205 | 3 | 0.07 | 0.05 | 0.28 | 0.60 | 0.10 | 0.05 | 0.04 | 0.85 |
| 2 | 481 | 4 | 0.06 | 0.06 | 3.71 | 0.05 | 0.03 | 0.06 | 0.68 | 0.41 |
| 3 | 478 | 4 | 0.03 | 0.13 | 2.74 | 0.10 | 0.01 | 0.13 | 3.01 | 0.08 |
| 4 | 392 | 4 | 0.04 | 0.05 | 0.11 | 0.74 | 0.07 | 0.06 | 5.04 | 0.03 |
| 5 | 219 | 3 | 0.08 | 0.05 | 0.00 | 0.95 | 0.06 | 0.08 | 2.01 | 0.16 |
| 6 | 492 | 4 | 0.02 | 0.13 | 7.35 | 0.01 | 0.01 | 0.10 | 6.66 | 0.01 |
| 7 | 592 | 5 | 0.12 | 0.05 | 0.19 | 0.67 | 0.05 | 0.07 | 0.30 | 0.58 |
| 8 | 416 | 4 | 0.05 | 0.09 | 1.62 | 0.20 | 0.03 | 0.08 | 0.85 | 0.36 |
| 9 | 165 | 3 | 0.12 | 0.05 | 0.08 | 0.77 | 0.18 | 0.09 | 0.40 | 0.53 |
| 10 | 387 | 4 | 0.05 | 0.06 | 0.25 | 0.62 | 0.05 | 0.07 | 0.03 | 0.86 |
| 11 | 598 | 5 | 0.07 | 0.05 | 0.21 | 0.64 | 0.05 | 0.05 | 1.11 | 0.29 |
| 12 | 198 | 3 | 0.03 | 0.05 | 1.01 | 0.31 | 0.07 | 0.08 | 0.39 | 0.53 |
| 13 | 355 | 4 | 0.10 | 0.04 | 0.03 | 0.86 | 0.10 | 0.04 | 0.27 | 0.61 |
| 14 | 444 | 4 | 0.05 | 0.07 | 3.52 | 0.06 | 0.06 | 0.06 | 4.42 | 0.04 |
| 15 | 462 | 4 | 0.04 | 0.10 | 4.16 | 0.04 | 0.03 | 0.11 | 1.03 | 0.31 |
| 16 | 336 | 4 | 0.13 | 0.05 | 0.41 | 0.52 | 0.09 | 0.03 | 1.28 | 0.26 |
| 17 | 453 | 4 | 0.03 | 0.08 | 3.49 | 0.06 | 0.05 | 0.06 | 1.87 | 0.17 |
| 18 | 413 | 4 | 0.07 | 0.06 | 0.58 | 0.45 | 0.08 | 0.07 | 0.03 | 0.87 |

number of degrees of freedom was equal to 1 (4-1-2). The observed and expected frequencies are given in Table 3 and 4.

The obtained $\chi2$-values as well as the right tail probabilities ($P\ \chi2$) are given in Table 2. For 5 of the 36 cases (18 holes for round 1 and 18 holes for round 2) $\chi2$ was significant at the 5%-level. However, in only one case $\chi2$ was significant at the 5%-level in both rounds. This was the case for hole 6. Therefore, if the null hypothesis would be accepted if at least in one of the two rounds $\chi2$ is not significant, then Hardy's model applies in the majority of the holes (17 out of 18).

## Discussion

The fact that Hardy's model applied to the majority of holes is a very satisfactory result. There was only one hole in which it did not apply, namely hole 6, which was denoted as a par four hole. A *par* is apocryphally described as an abbreviation for "professional average result". According to this description one might expect, that, to be denoted as, for example, a par four the average score of a hole for professionals should be between the values 3.5 and 4.5. According to this description

**Table 3 Score Frequencies for Round 1 (*N* = 156, -1: a Birdie or Less, 0: a Par, 1: a Bogey and 2: a Double Bogey or More).**

| | | Observed | | | | Expected | | | |
|---|---|---|---|---|---|---|---|---|---|
| **Hole** | **Par** | **-1** | **0** | **1** | **2** | **-1** | **0** | **1** | **2** |
| 1 | 3 | 21 | 117 | 17 | 1 | 21.0 | 117.7 | 15.8 | 1.5 |
| 2 | 4 | 21 | 104 | 30 | 1 | 21.5 | 106.8 | 23.9 | 3.8 |
| 3 | 4 | 11 | 79 | 51 | 15 | 10.7 | 85.9 | 42.1 | 17.3 |
| 4 | 4 | 18 | 115 | 20 | 3 | 17.7 | 114.3 | 21.2 | 2.7 |
| 5 | 3 | 22 | 114 | 18 | 2 | 22.0 | 114.2 | 17.8 | 2.0 |
| 6 | 4 | 7 | 76 | 59 | 14 | 6.8 | 84.7 | 44.5 | 19.9 |
| 7 | 5 | 56 | 80 | 18 | 2 | 55.0 | 80.9 | 17.4 | 2.6 |
| 8 | 4 | 19 | 91 | 39 | 7 | 18.5 | 95.3 | 33.2 | 8.9 |
| 9 | 3 | 33 | 106 | 15 | 2 | 33.0 | 105.8 | 15.5 | 1.7 |
| 10 | 4 | 18 | 109 | 24 | 5 | 17.7 | 107.5 | 26.2 | 4.6 |
| 11 | 5 | 38 | 91 | 23 | 4 | 35.6 | 93.1 | 23.2 | 4.0 |
| 12 | 3 | 9 | 124 | 22 | 1 | 9.0 | 125.4 | 19.4 | 2.2 |
| 13 | 4 | 41 | 97 | 16 | 2 | 40.3 | 97.7 | 16.1 | 1.8 |
| 14 | 4 | 20 | 97 | 36 | 3 | 20.5 | 100.6 | 28.8 | 6.1 |
| 15 | 4 | 14 | 87 | 47 | 8 | 13.7 | 93.1 | 37.3 | 11.9 |
| 16 | 4 | 49 | 88 | 16 | 3 | 46.5 | 89.4 | 17.6 | 2.5 |
| 17 | 4 | 9 | 103 | 40 | 4 | 9.9 | 106.6 | 32.3 | 7.2 |
| 18 | 4 | 29 | 98 | 26 | 3 | 28.1 | 98.9 | 24.5 | 4.4 |

**Table 4 Score Frequencies for Round 2 (*N* = 156, -1: a Birdie or Less, 0: a Par, 1: a Bogey and 2: a Double Bogey or More).**

| | | Observed | | | | Expected | | | |
|---|---|---|---|---|---|---|---|---|---|
| Hole | Par | -1 | 0 | 1 | 2 | -1 | 0 | 1 | 2 |
| 1 | 3 | 28 | 110 | 16 | 2 | 28.0 | 109.9 | 16.3 | 1.8 |
| 2 | 4 | 11 | 113 | 29 | 3 | 10.9 | 114.1 | 26.6 | 4.4 |
| 3 | 4 | 4 | 80 | 55 | 17 | 3.9 | 86.7 | 45.4 | 20.0 |
| 4 | 4 | 29 | 104 | 16 | 7 | 28.2 | 101.0 | 23.0 | 3.8 |
| 5 | 3 | 16 | 105 | 32 | 3 | 16.0 | 108.0 | 26.8 | 5.2 |
| 6 | 4 | 5 | 90 | 53 | 8 | 4.9 | 98.1 | 40.1 | 12.9 |
| 7 | 5 | 27 | 95 | 27 | 7 | 25.8 | 94.0 | 29.5 | 6.7 |
| 8 | 4 | 12 | 106 | 29 | 9 | 11.9 | 102.0 | 33.7 | 8.4 |
| 9 | 3 | 45 | 87 | 19 | 5 | 45.0 | 85.7 | 21.1 | 4.1 |
| 10 | 4 | 21 | 103 | 27 | 5 | 20.5 | 104.1 | 26.5 | 4.9 |
| 11 | 5 | 27 | 101 | 26 | 2 | 25.8 | 103.2 | 23.3 | 3.6 |
| 12 | 3 | 20 | 104 | 28 | 4 | 20.0 | 105.6 | 25.6 | 4.8 |
| 13 | 4 | 42 | 98 | 15 | 1 | 40.3 | 100.2 | 14.1 | 1.4 |
| 14 | 4 | 25 | 99 | 31 | 1 | 24.3 | 103.4 | 24.2 | 4.1 |
| 15 | 4 | 10 | 89 | 45 | 12 | 9.8 | 92.6 | 39.8 | 13.8 |
| 16 | 4 | 36 | 106 | 14 | 0 | 34.8 | 108.1 | 12.1 | 0.9 |
| 17 | 4 | 21 | 109 | 20 | 6 | 20.6 | 105.4 | 25.5 | 4.5 |
| 18 | 4 | 31 | 93 | 26 | 6 | 30.9 | 92.8 | 26.6 | 5.7 |

the par of all holes were denoted correctly except for hole 6 with an average of 4.54 during round 1. During round 2 the average was equal to 4.41, which is close to 4.5. In comparison with all of the par four holes, hole 6 has the largest length (492 yards). The Hardy distribution appears to be sensitive to aberrations of the par denotation of a hole. The average score values of hole 6 for round 1 and round 2 are close to 4.5 indicating, that this hole actually must be considered as a kind of hybrid between a par four and a par five hole. This might be the main reason that a par four Hardy distribution does not hold for this hole.

In an article titled "G.H. Hardy's Golfing Adventure" R. B. Minton (Minton, 2010) discussed the Hardy model for the case $p$=$q$. In this special case one would expect, for each par, a positive product moment correlation between the estimates of $p$ and $q$ across holes and days. Since the number of data points was very small for the par three and par five holes, the correlation between $\hat{p}$ and $\hat{q}$ for only the par four holes was computed. Not a positive, but a significant negative correlation was found ($r$=-0.730, $p$ = .000, $N$ = 24). The negative correlation between $\hat{p}$ and $\hat{q}$ could be explained as follows. It is natural to assume, that the probabilities $p$ and $q$ are each dependent on the *skill* of the player and the *difficulty* of the hole. However, since the assumption is made, that all players are equivalent, for each par the differences between the various $p$'s and $q$'s across holes are only dependent on the difficulty of the holes. For holes of the same par, such as a par four, the difficulty of the holes will be mainly dependent on the length ($l$) of the hole. This would mean, that a negative correlation should be found between the length of the hole and $\hat{p}$ and a positive correlation between the length of the hole and $\hat{q}$. This is exactly what has been found ($r_{\hat{p}l} = -0.822$, $p = .000$; $N = 24$; $r_{\hat{q}l} = +0.736$, $p = .000$, $N = 24$). The partial correlation $r_{xy;z}$ measures the strength of a relationship between two variables $x$ and $y$ while controlling the effect of another variable $z$. The partial correlation between $\hat{p}$ and $\hat{q}$ (partialling out the length of a hole) was not significant ($r_{\hat{p}\hat{q};l} = -0.325$, $df = 21$, $p = .131$)

One may wonder what practical use there would be in knowing the theoretical distribution of golf hole scores. Hardy once commented "Nothing I have ever done is of the slightest practical use." (see Wikipedia on G. H. Hardy). While Hardy's model is an oversimplication of golfing results, it nevertheless satisfactorily explains the shape of the frequency distribution of golf hole scores. That the observed frequency distribution differed significantly from the expected frequency distribution may be a sign that something has gone wrong in the par designation of the hole. It has been shown that there is a relation between the parameters $p$ and $q$ on the one hand and the length of the hole on the other hand. In a separate study (van der Ven, unpublished), using the results of the KLM Open 2012, it was found that for par three holes there exists a relationship between the parameters $p$ and $q$ and the latitude of the green. More of such relations between the parameters $p$ and $q$ and possible measures of the difficulty of the hole may be found in the future. This may ultimately make it possible to predict the theoretical frequency distribution of the hole scores before the golf tournament even takes place. Such information could be of critical importance to the organizers as well as to the participants of a tournament. For the organizers the shape of the distribution may give rise to changing the hole's overall layout. For the participants the shape of the distribution gives information about the probabilities of the hole scores or, stated differently, which scores are to be expected for that particular hole.

# Notes

1. Shot value is a measure of the quality of a shot relative to a scratch golfer's average shot from a given situation. (see Broadie, 2008, first line of page 3)

2. The details of the algorithms used by Maple's core functions (including 'fsolve') are proprietary and therefore, specific details on their inner workings are not available to the general public.

3. Maplesoft, a division of Waterloo Maple Inc., Waterloo, Ontario.

## Acknowledgments

The author would like to thank Bruce Jenks and Mark Broadie for their valuable comments.

# References

Broadie, M.Crews, D., & Lutz, R.Editors. (2008). Assessing golfer performance using Golfmetrics. Chapter 34 in *Science and Golf V: Proceedings of the 2008 World*

*Scientific Congress of Golf*, Editors: D. Crews and R. Lutz, Energy in Motion Inc., Mesa, Arizona, 253-262.

Clark, R.D., III. (2004). On the Independence of Golf Scores for Professional Golfers. *Perceptual and Motor Skills, 98*(2), 675–681. PubMed doi:10.2466/pms.98.2.675-681

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences*. New York: Academic Press.

Cohen, G.L. (2002). On a theorem of G. H. Hardy concerning golf. *The Mathematical Gazette, 86,* 120–124. doi:10.2307/3621594

Edgell, E., & Noon, S.M. (1984). Effect of violation of normality on the t-test of the correlation coefficient. *Psychological Bulletin, 95,* 576–583. doi:10.1037/0033-2909.95.3.576

Hardy, G.H. (1945). A mathematical theorem about golf. *The Mathematical Gazette, 29,* 227. doi:10.2307/3609265

Lord, F.M., & Novick, M.R. (1968). *Statistical Theories of Mental Test Scores*. New York: Addison Wesley.

Minton, R.M. (2010). G. H. Hardy's golfing adventure. *Mathematics and Sports*, Joseph A. Gallian, ed. MAA, 169 - 179.

van der Ven, A.H.G.S. (2012). The Hardy distribution for golf hole scores. *The Mathematical Gazette, 96,* 428–438.