A Statistical Analysis of 10-pin Bowling Scores and an Examination of the Fairness of

Alternative Handicapping Systems[1]

Sarah Keogh and Donal O'Neill[2]

Abstract

Using data on approximately 1240 games of bowling we examine the statistical properties of 10-pin bowling scores. We find evidence that that the distribution of bowling scores is approximately log-normally distributed with a common variance across players. This allows us to consider the effectiveness of alternative handicapping systems in allowing less skilled bowlers to compete against more skilled opponents. We show that the current system mitigates against bowlers of low skill and propose a new system which we show works well in equalising the playing field across all match-ups.

I: Introduction

Primitive forms of bowling can be traced back as far as 3200BC (www.tenpinbowling.org 2011). Ten pin bowling as we know it, is thought to have originated from the German nine pin bowling game called Kegal (www.tenpinbowling.org 2011). Today, it is estimated that the sport of ten pin bowling is played by over 100 million people in more than 90 countries, making it one of the most popular games played in the world.[3] With this in mind, it's surprising how little quantitative analysis has been conducted on the sport.

Previous research has examined the correlation in scores between balls rolled Neal (2004) using a binomial distribution while Dorsey-Palmateer and Smith (2004) used empirical data to examine the phenomenon of a "hot-hand" whereby the probability of a strike with a given ball depends on the score obtained with the previous ball. There has also been a small body of work looking at the distribution of bowling scores (Cooper and Kennedy (1990) and Hohn (2009)) that did not use observed data. The only paper we are aware of that uses empirical data to analyse the distribution of bowling scores is work carried out by Chen and Swartz (1994), which used data on five pin bowling scores, a less-popular variant of ten-pin bowling played only in Canada. As noted by the authors however, their results cannot be applied to ten pin bowling due to differences in rules, equipment and scoring.

In this paper we carry out a statistical analysis of scores from approximately 1024 games of ten pin bowling to determine the distribution of these scores. Since the score of a player in a single score is the sum over a number of frames one might expect something approximating normality if a central limit theorem is in operation. We examine normality of

---

[3] See World Tenpin Bowling Association website http://www.worldtenpinbowling.com/

both the raw scores and a suitably chosen transformation of these scores. We then use the results of this analysis to evaluate the effectiveness of alternative handicapping systems currently used. The latter is very topical among elite bowlers because if a bowler believes a certain handicapping system to be unfair towards them, they will not enter leagues or competitions which use that system. Section II introduces the game of ten pin bowling and describes the data set used in the analysis. The statistical analysis is conducted in section III, while section IV compares current handicapping systems and proposes a new, alternative system.

II: The Game and the Data Set

Ten pin bowling is a competitive sport in which a "bowler" rolls a bowling ball down the lane with the aim of scoring points by knocking down as many pins as possible. Ten pin bowling balls have a smooth surface bar two finger holes and a thumb hole. They have a maximum diameter of 8.59 in., and range in weight from 6lb. to 16lb. The ball is rolled down a wooden or synthetic lane with the objective of knocking down pins. The pins themselves are 4.75 in. wide at their widest point and 15 in. tall. They weigh 3 lb., 6 oz. The bowling lane itself is 60 ft. from the foul line to the centre of the headpin. Today it is estimated that over 100 million bowlers play the sport in over 90 countries, making it one of the most popular participant sport in the world.

A game of ten pin bowling consists of 10 frames. A bowler's score is the sum of points over the game. Frames 1 to 9 consist of a maximum of 2 legally delivered balls rolled by the same bowler on the same lane. Frame 10 consists of a maximum of three legally delivered balls by the same bowler on the same lane. In a frame of bowling, if a bowler knocks down all ten pins with their first ball, then a strike is awarded. If all ten pins are

knocked down using two balls in any frame, a spare is awarded. A strike is worth 10 points, plus the points received from the next two balls rolled. A spare is worth 10 points plus the points received from the next ball rolled. A bowler who bowls a strike in the 10th frame is awarded two extra balls, which allow the awarding of bonus points. A bowler who bowls a spare in the 10th frame is awarded one additional ball to allow for the bonus points. The maximum score in ten pin bowling is 300, and is achieved by having 12 strikes in a row in one game. The typical average score for a woman is 153, and the typical average score for a man is 173.[4]

The data set used in this paper was obtained from an adult ten pin bowling league in Alsaa Bowling Centre, Dublin Airport, Ireland. The league is 18 weeks in duration. There are 23 bowlers in this league. These 23 bowlers played 3 games every Monday evening for the 18 weeks. This gives a maximum of 54 games per bowler with possibly some missing values. This results in approximately 1240 games. Bowling scores vary in the data set. The maximum score is 289, bowled by player 20, and the minimum score is 114 bowled by player 16. Skill levels among the 23 bowlers differ. The average scores range from 153 (player 18) to 220 (player 7). Further summary statistics are given in Table 1.

III: Analysis

    a. Raw Scores:

The first stage in this analysis will be to investigate if the raw data is approximately normally distributed. To begin we construct kernel density estimates of the underlying density. Kernel density estimation is a non-parametric way of estimating the probability density function of a random variable (Sheather 2004). The kernel density estimates for each of the 23 bowlers are

---

[4] http://www.ten-pinbowling.com/faq.htm#3

given in Figure 1. In these figures a normal density with mean and variance derived from the raw data is also provided for comparison. These graphs show deviations from normality for a number of players For instance results for players 1, 2, 13 and 20 seem to be skewed to the right, while players 5, 21 and 23 seem to have fatter tails that the underlying normal distribution.

We supplement this graphical approach with a number of tests. The tests used are the Shapiro-Francia test (Shapiro and Francia (1972), Royston (1983)) and the Skewness Kurtosis test (D'Agostino et al (1990), Royston (1991)). The Shapiro-Francia test equals the squared correlation between the ordered sample values and the expected order statistics of a standard normal and thus measures the straightness of the normal probability plot of the sample values. The Skewness Kurtosis test presents a test for normality based on skewness and another based on kurtosis and then combines the two tests into an overall test statistic. Each test is conducted for each player in the data set. Then the p values are combined using Fisher's Method (Fisher 1970). Fisher's method combines p values for each test into one test statistic using the formula:

$$P = -2\sum_{i=1}^{k} \ln(p_i)$$

where p_i is the p value for the ith hypothesis test. When all the null hypotheses are true, and the p_i are independent, P has a chi-squared distribution with $2k$ degrees of freedom (Fisher 1925). The results for both tests are presented in Tables 2 and 3. For both tests we reject the null hypothesis that all game scores are normally distributed.

b. Transformed Scores

Having rejected normality of the raw scores we now consider whether it is possible to find a transformation of the underlying scores which is well represented by a normal distribution. To do this we consider Box-Transformations of the scores (Box and Cox 1964, Spitzer 1980). The original form of the Box-Cox transformation, as appeared in their 1964 paper, takes the following form:

$$y^{(\lambda)} = \begin{cases} \dfrac{(y^{\lambda} - 1)}{\lambda} & \lambda \neq 0 \\ ln(y) & \lambda = 0 \end{cases}$$

Formally we posit the existence of a λ such that bowling scores for player i in game j, y_{ij}, can be written as :

$$y_{ij}^{(\lambda)} = \beta_i + \varepsilon_{ij}$$

where ϵ_{ij} is a draw from a normal distribution with mean zero and variance σ_i^2. Under the assumption of normality the density function for the jth observation of player i, can be written

$$f(\varepsilon_{ij}) = (2\pi\sigma_i^2)^{-1/2} exp(-\varepsilon_{ij}^2 / 2\sigma_i^2)$$

The likelihood contribution for player i can then be written in terms of observed bowling scores y_{ij} :

$$L_i^* = \prod_j f(\varepsilon_{ij}) y_{ij}^{\lambda-1}$$

Where the last term is the Jacobian of the transformation from ϵ to y. Taking logs we get

$$ln L_i^* = c - T_i / 2 ln(\sigma_i^2) - \left[(2\sigma_i^2)^{-1} \sum_{j=1}^{T_i} (y_{ij}^{(\lambda)} - \beta_i)^2 \right] + (\lambda - 1) \sum_{j=1}^{T_i} ln(y_{ij})$$

6

where c is a constant. The log likelihood function across all players is then:

$$ln\, L^* = \sum_{i=1}^{N} ln\, L_i^*$$

Since β_i is just the mean of $y_{ij}^{(\lambda)}$, we can concentrate it out of the likelihood function by replacing it with its maximum likelihood estimate

$$\hat{\beta}_i = \left[(T_i)^{-1} \sum_{j=1}^{T_i} \left(y_{ij}^{(\lambda)} \right) \right] = \overline{y}_i^{(\lambda)}$$

We can simplify this further by solving for the parameter σ_i^2 directly from its first order conditions in terms of the other parameters and the data. The maximum likelihood estimate of σ_i^2 is given by :

$$\hat{\sigma}_i^2 = \left[(T_i)^{-1} \sum_{j=1}^{T_i} \left(y_{ij}^{(\lambda)} - \overline{y}_i^{(\lambda)} \right)^2 \right]$$

Substituting this expression into the individual likelihood function we get a final concentrated likelihood function depending only on λ and data.

$$ln\, L_i^* = \tilde{c} - T_i / 2 \left[ln \left(\frac{\sum_{j=1}^{T_i} \left(y_{ij}^{(\lambda)} - \overline{y}_i^{(\lambda)} \right)^2}{T_i} \right) \right] + (\lambda - 1) \sum_{j=1}^{T_i} ln\left(y_{ij} \right)$$

The concentrated log-likelihood across all players can then be written as

$$ln\, L^* = \sum_{i=1}^{N} \left[-T_i / 2 \left[ln \left(\frac{\sum_{j=1}^{T_i} \left(y_{ij}^{(\lambda)} - \overline{y}_i^{(\lambda)} \right)^2}{T_i} \right) \right] + (\lambda - 1) \sum_{j=1}^{T_i} ln\left(y_{ij} \right) \right]$$

Maximising the concentrated likelihood function with respect to λ provides an estimated value $\hat{\lambda}$ =-.148 with a standard error of .199 (lnL=-3331.2102). Since we cannot reject λ=0 (i.e the log transformation) we adopt this for convenience throughout the remainder of the paper.

Table 4 gives a brief description of the transformed log bowling scores. The kernel density graphs for the transformed bowling scores of each of the 23 bowlers are given in Figure 2 and the results of the formal tests are given in Tables 5 and 6. Looking at the Shapiro-Francia test results in Table 5 we see that while individually normality is rejected for 2 out of the 23 players, Fisher's combined test fails to reject the null-hypothesis of normality across all players. Table 6 presents the results of the Skewness Kurtosis. At the player level normality is rejected for only one of the twenty three players. Perhaps somewhat surprisingly Fisher's combined test for normality across all players is rejected, the result driven by one single player. When this one player is omitted from the analysis the combined F-test fails to reject normality across the remaining players.

Analysis of the handicapping system will require comparing scores of high and low ability bowlers. To do this it is useful to first examine whether or not the variance of bowling scores depends on a player's ability. Figure 3 plots the variance of transformed scores against average score for each of the twenty three players, as well as the predicted variances from a least squares regression through the data.[5] There is little evidence of a systematic relationship between variability and average ability.[6] The estimated slope coefficient of the linear regression is -.002 with a robust standard error of .012.

---

[5] The labels on each observation indicate the number of games played by the player in question.

[6] In contrast the raw scores exhibit a significant positive relationship between ability and variability. This may reflect the dramatic impact of strikes and spares (most likely to be bowled by high ability bowlers) on scores in

IV: Handicapping

In bowling leagues the range of abilities can be wide. In our sample the average scores ranged from 155 to 220. In light of this a league may adopt a handicapping system. A handicap is simply a number added to a player's gross score after each game which is dependent on the player's average ability. Handicaps are used to allow less skilled bowlers compete with highly skilled bowlers on a more equal playing field. The idea is the bowler with the higher average should always have a smaller advantage given the handicapping system. This provides an incentive for bowlers to improve their bowling scores. The advantage of the higher average player should not be so great that it discourages the lower average player. Formally the US Bowling Congress defines a handicap league as follows

"A handicap league is one in which handicap is added to a bowler's score to place bowlers and teams with varying degrees of skill on as equitable a basis as possible for scheduled competition." (USBC 2011)

However, they are not explicit as to what is meant by equitable. The Council of National Golf unions, the association charged with determining the handicap scores for golfers in Britain in Ireland, are more explicit in stating that "A golf handicap allows players of all levels of golfing ability to compete against each other on a fair and equal basis."[7] Use of the term *equal* suggests that the purpose of a handicap system is to equalise the probability of victory across ability levels. It is this feature of the handicap system that we examine in the context of 10-pin bowling.

---

a given frame. This contrasts to other sports such as golf where there is some evidence that variability declines with ability (Bingham and Swartz (2000)).

[7] http://www.congu.com/welcome.htm

There have been some studies of the fairness of handicapping system in other sports (see McHale (2010) and Bingham and Swartz (2000) for examples in golf). However there has been little empirical analysis on bowling handicaps. Chen and Swartz (1994) carried out a statistical analysis of 5-pin bowling, a variant of bowling, which is played only in Canada. However, the scoring in this version of the game differs significantly from the more popular 10-pin bowling version and little attention was paid to developing a more equitable handicapping system.

There are several handicapping systems used in 10-pin bowling and they can vary from league to league. When considering fairness in this context we need to consider both players. It is important the handicapping system is fair to the lower average player in order to provide an incentive to increase their bowling scores. However, we want to avoid systems that are so generous such that the lower average bowler can win without playing very well; the incentives from such a system are wrong for both players. Since the low ability player can win without playing very well they have little incentive to improve their game. This in turn may prohibit such players from competing in leagues that don't use handicaps (scratch leagues). On the other hand if the high ability player believes the system is such that they will lose even when playing well, they will have little incentive to try hard and may opt out of leagues using the system.

At present there are a variety of used in Ireland. Depending on which of the systems is used the final score of a player with average ability m_i (determined over previous games played by the player) is adjusted by adding a bonus to their raw score. The bonus is determined by a weighting factor (w) and a par or scratch score P as follows:

$$\text{Bonus} = w(P - m_i)$$

The most common weighting factor is 80% or .8, though weighting factors as low as .66 are used. Typical scratch scores are 220 or 200. Cleary the higher the weighting factor and par score the more generous the system is. Typical handicap systems use w=.8 and P=220. Under this system a player with an average of 120 would be credited with an extra 80 pins (or equivalently have an extra 80 points added to their raw score) under this system. Reducing w to .66 and P to 200 would result in the same player only receiving only an additional 52 points. Although the base or scratch score used to calculate the handicap affects the number of pins added to a player's score, the choice of base will not affect the probability of victory provided the scores of both players are adjusted under the handicap. So for instance in cases where negative handicaps are allowed (whereby players above the base score have pins deducted from their gross score; "plus handicappers" in golfing terms) then only the weighting factor will impact on the probability of victory.

With w=.8 and P=220 a bowler with an average of 160 would have to beat a lower ability opponent with an average say of 120 by more than 32 (.8*40) points in order to win the game. Letting the higher ability players score be denoted by X and that of the lower ability player by Y, then to establish the probability that the favourite (high ability player) wins, we need to determine Pr(X>Y+32). More generally we can think of variations in handicap systems by considering Pr(X>Y+h) for a range of h (corresponding to different handicap systems). Letting f(X,Y) denote the joint distribution of X and Y, we can write this probability as :

$$Pr(X > Y + h) = \int_h^\infty \int_0^{x-h} f(X,Y)\,dxdy \qquad (1)$$

If we further assume that X and Y are independent then using theory on the distribution of functions of random variables, together with $Z = ln(x) \sim N(\mu_x, \sigma_x^2)$ and

$W = ln(y) \sim N(\mu_y, \sigma_y^2)$ we can derive f(X,Y) as[8] :

$$f(X,Y) = -\frac{1}{2\sigma^2} exp\left[\left\langle ln(x) - \left\{ ln(\bar{x}) - \frac{\sigma^2}{2} \right\} \right\rangle^2 + \left\langle ln(y) - \left\{ ln(\bar{y}) - \frac{\sigma^2}{2} \right\} \right\rangle^2 \right] \frac{1}{x} \frac{1}{y}$$

Evaluation of the subsequent integral in (1) requires either numerical methods or some additional approximations (Chen and Swartz (1994). Instead we use Monte-Carlo techniques to evaluate these probabilities under two handicap systems; corresponding to weighting factors of .8 and .66 respectively. We compare the win probabilities of the favourite using raw and gross scores under both handicap systems.

To do this we simulate bowling scores for a given player with average ability m_i by generating random draws from a $N\left(\log(m_i) - \frac{\sigma^2}{2}, \sigma^2 \right)$ distribution. We use the pooled variance across all players which is estimated as $\hat{\sigma}_p^2 = .01957$.[9] We then take the exponential of the random draw as our raw bowling score. We adjust the score by applying each of the handicap systems and then compare outcomes across players of different ability. We choose 7 values of m_i from 100 to 220 in intervals of 20. When $m_i \leq m_j$, we refer to player i as the underdog and player j as the favourite. For each pair of players, 10,000 bowling scores are generated and adjusted as described above. We then estimate the probability of the favourite defeating the

---

[8] Here we have used the further simplification that $\sigma_x^2 = \sigma_y^2 = \sigma^2$.

[9] This pooled variance is almost identical to the predicted variance across all ability levels given by the fitted regression shown in Figure 3. Using the fitted line the estimated variance ranges from .0208 for the lowest ability to .0199 for the highest ability.

underdog by computing the fraction of the 10,000 games the favourite has won based on raw and adjusted scores.

Table 7 details the probabilities of the favourite winning across each possible match-up using both raw unadjusted scores (first entry in each cell), adjusted according to most generous system (2nd entry in each cell) and adjusted according to last generous system (3rd entry in each cell). Looking first at the probabilities based on raw scores we see that even with modest differences in ability the probability of the favourite winning is consistently over 75% and can quickly rise to above 90%. It is this imbalance in the probability of winning that motivates the use of handicap systems. The 2nd and 3rd entry in each cell reports the probability of the favourite winning using the scoring methods in the most generous and least generous handicap systems. As expected the adjustment reduces the favourites advantage, however the playing field remains far from level. The probability of a player with ability level 180 beating a player with ability level 100 ranges from 70%-83% depending on the system, despite the difference in skills having been taken into account in the handicap system. It is clear from these results that even under the most generous handicap system currently used in Irish bowling leagues the system remains inequitable with the more able players more likely to win, especially when the ability differential is relatively large.

These simulated probabilities are based on our estimated variance and the assumption of log normality. It is interesting to compare how these probabilities compare to the actual outcomes in the bowling league under analysis. As well as the individual scores we also have access to the average score of each player in the previous 20 games bowled (which is the basis for the handicap score). For example at the start of the league player 7 had an average of 220 and player 15 an average of 160. Since this was a team league these players never competed directly against each other in head to head games. Nevertheless both players

bowled in all 54 games, with their games taking place on the same evenings, at the same time and in the same bowling alley. One check of our results is to compare the scores recorded by both of these players on games played at the same time as if they were competing against each other and to determine the victor in each of these 54 games. Using just the raw scores we find that the high ability player would have emerged victorious in 53 of the 54 games. The one exception was in game 49 when player 15 scored 189 compared to player 7's score of 174. The higher ability player won all the remaining games with a margin of victory ranging from 9 to 121 points. This comparison yields an estimated probability of victory for the high ability player equal to .98, which compares favourably to our predicted estimate of .95. We then repeat the analysis this time adjusting scores using a par score of 220 and a weighting factor of .8 (similar to what is actually used in the league). With these adjusted scores the number of games won by the favourite falls to 35, giving a probability of victory equal to .648. This again compares very well to our simulated probability of .621. It would thus appear that our simulated probabilities based on 10,000 simulated games provide estimates that match very well to recorded victories, in actual league play based on much fewer games.

We now examine what changes, if any, to the handicap system would produce more equitable outcomes. As noted earlier the key feature of handicap system determining the probability of victory is the weighting factor. The United States Bowling Congress (2011) recommends that their leagues adopt a weighting of 100% in order to equalise probabilities. Table 8 shows the outcome probabilities under such a system. As expected this system produces a much more level playing field with probabilities closer to .5. However, under this scheme the odds of victory shift towards the lesser ability player. In all cases the probability of the favourite wining is less than .5. Some authors (McHale 2010) argue that such a system would distort incentives. Thus given the observed distribution of bowling scores in Ireland

we would propose a weighting factor slightly less than 1. Table 9 shows the results obtained when the weighting factor is .95. We see that across all match-ups the probability of the favourite winning falls in the range .505-.547. This scheme thus results in a significant levelling of the playing field without removing the incentive for players to improve their skill level.

V: Conclusions

This paper examines the distribution of 10-pin bowling scores and uses the statistical results to examine the impact of alternative handicapped scoring systems. We show that while the raw bowling scores are not normally distributed, there is strong evidence that bowling scores are normally distributed after a log-transformation. The simplicity of this transformation helps when calculating the likelihood of victory in any match-up. Handicap scoring systems are used in many bowling leagues with the explicit purpose of allowing players to compete on a level playing field with players of lesser or greater skill. We use our results to examine the performance of currently used handicap systems in Ireland. We show that the current systems still leave large biases in favour of high ability players. Using a 100% weighting factor produces a more level playing field but goes too far in the sense that players have no incentive to improve their handicap. We propose a compromise weighting factor equal to .95 and shows that this system allows lower ability players to compete across all match-ups yet still provides some incentives for players to improve their game.

References

Bingham, D. and T. Swartz (2000), "Equitable Handicapping in Golf," *The American Statistician*, Vol. 54, no. 3, pp. 170-177.

Box, G. E. P., and Cox, D. R. (1964), "An analysis of Transformations", *Journal of the Royal Statistical Society,* 26, 211-243.

Chen, W. And Swartz, T. (1994), "Quantitative Aspects of Five-Pin Bowling", *The American Statistician,* 48, 92-98.

Cooper, C. N. and Kennedy, R. (1990), "A Generating Function for the Distribution of the scores of all possible Bowling Games", *Mathematical Magazine,* 63, 239-243.

D'Agostino, R. B., Belanger, A. and D'Agostino Jr. R. B. (1990), "A Suggestion for Using Powerful and Informative Tests of Normality, *The American Statistician,* 44, 316-321.

Dorsey-Palmeteer, R. And Smith, G. (2004), "Bowler's Hot Hand", *The American Statistician*, 58, 38-45.

Fisher, R.A. (1970), Statistical Methods for Research Workers 14th edition Hafner Publishing Company, New York

Hohn, J. L. (2009), "Generalized Probabilistic Bowling Distributions", *Unpublished Master's Thesis,* Western Kentucky University, Dept. of Mathematics.

McHale, I (2010), "Assessing the Fairness of the golf Handicap System in the UK," *Journal of Sports Sciences,* 28(10), pp. 1033-1041.

Neal, D. K. (2004), "Random Bowling", *The Journal of Recreational Mathematics,* 32, 125-129.

Royston, J. P. (1983), "A simple method for Evaluating the Shapiro-Francia W' Test of Non-Normality", *Statistician,* 32, 297-300.

Royston, J. P. (1991), "Comment on sg3.4 and an Improved D'Agostino Test", *Stata Technical Bulletin*, 3, 13-24.

Shapiro, S. S. And Francia, R. S. (1972), "An Approximate Analysis of Variance Test for Normality", *Journal of the American Statistical Association,* 67, 215-216.

Spitzer, J. J. (1982), "A Primer on Box-Cox Estimation", *The Review of Economics and Statistics*, 64 (2), 307-313.

United States Bowling Congress (2011) USBC Playing Rules and commonly Asked Questions http://usbcongress.http.internapcdn.net/usbcongress/bowl/rules/pdfs/2011-2012PlayingRules.pdf

Figure 1: Kernel density of raw bowling scores for each player
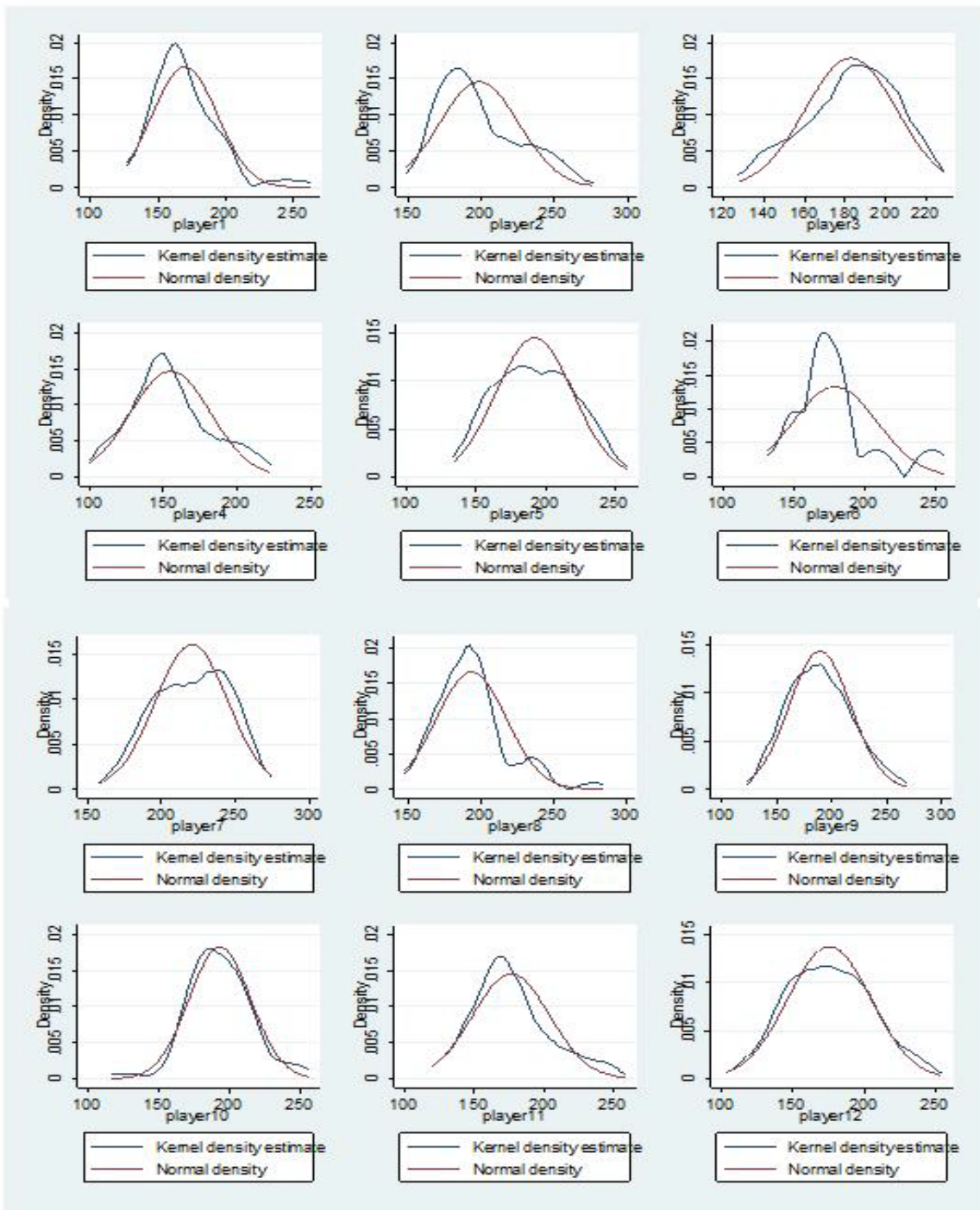
Figure 1 (continued)
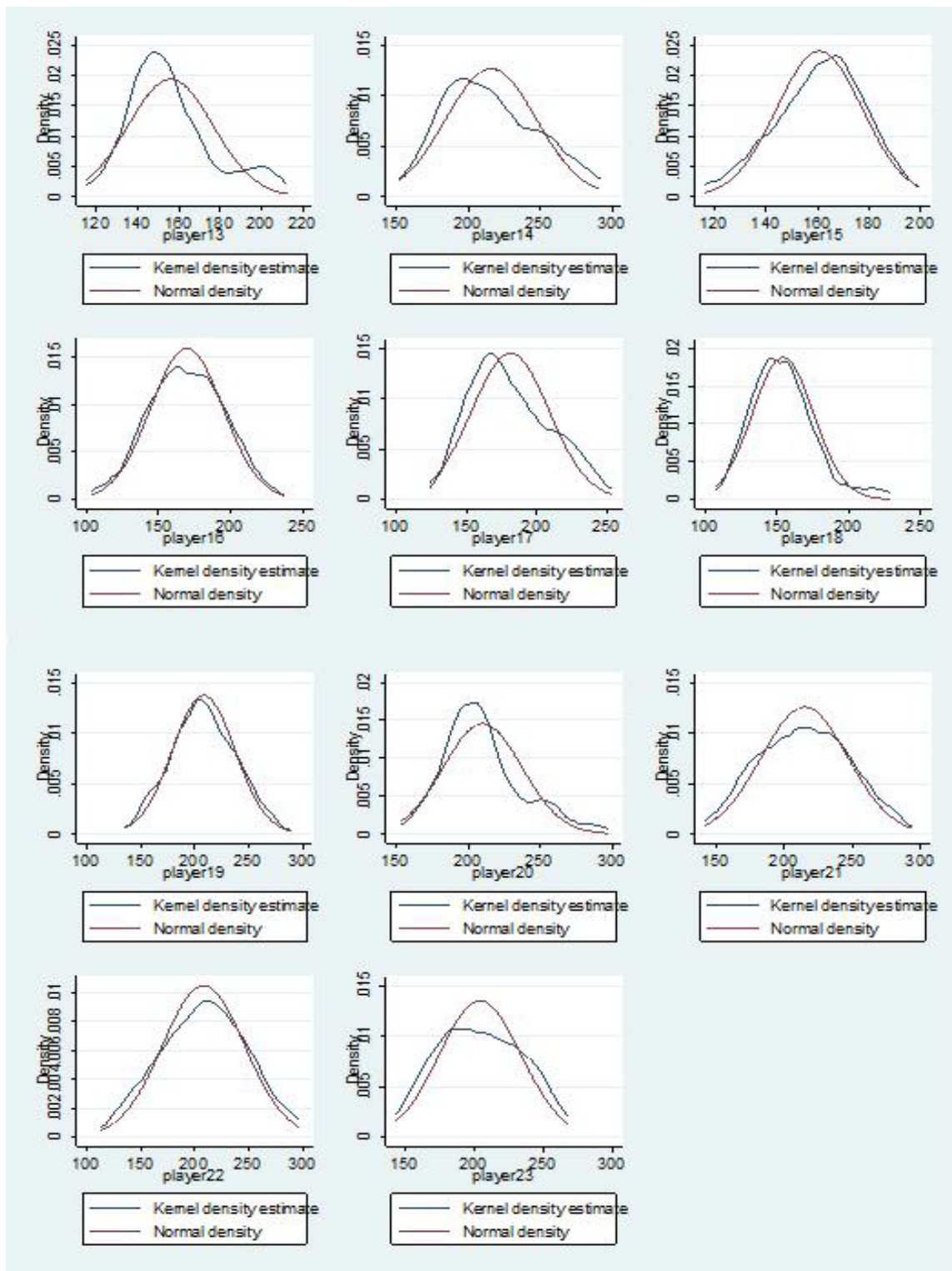
Figure 2: Kernel density of log transformed bowling scores for each player
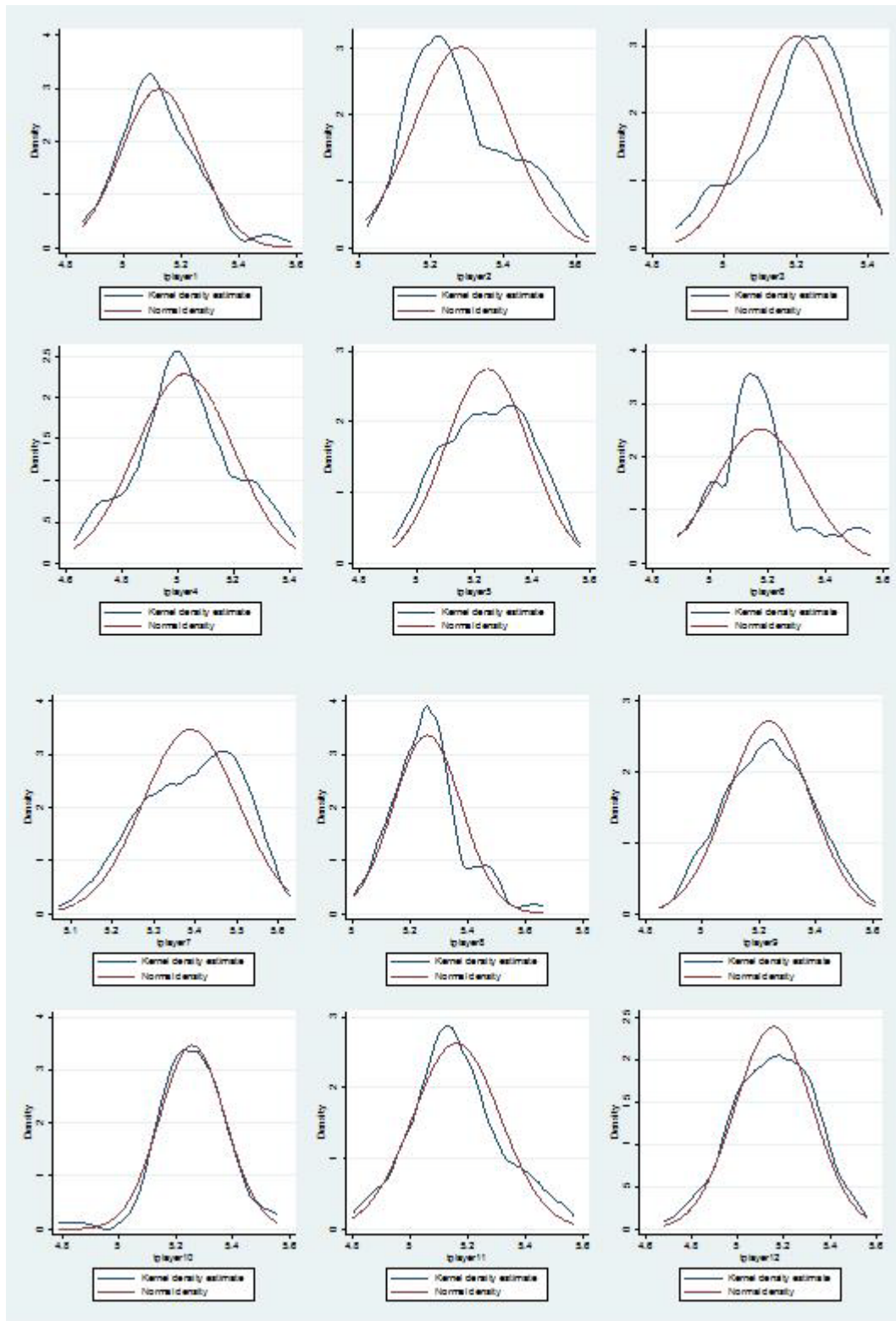
Figure 2: Continued
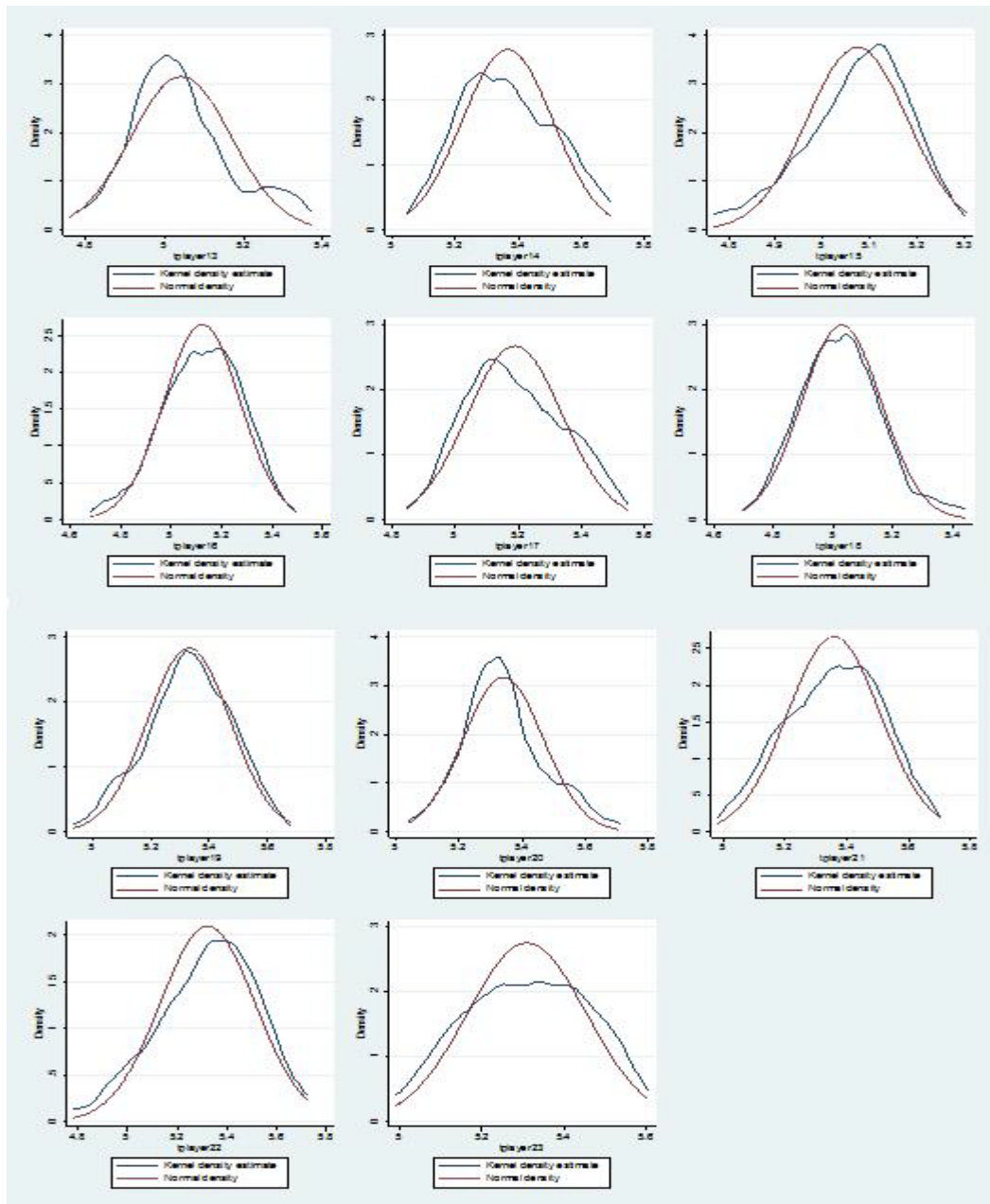
Figure 3: Variance of log transformed bowling scores and average bowling score
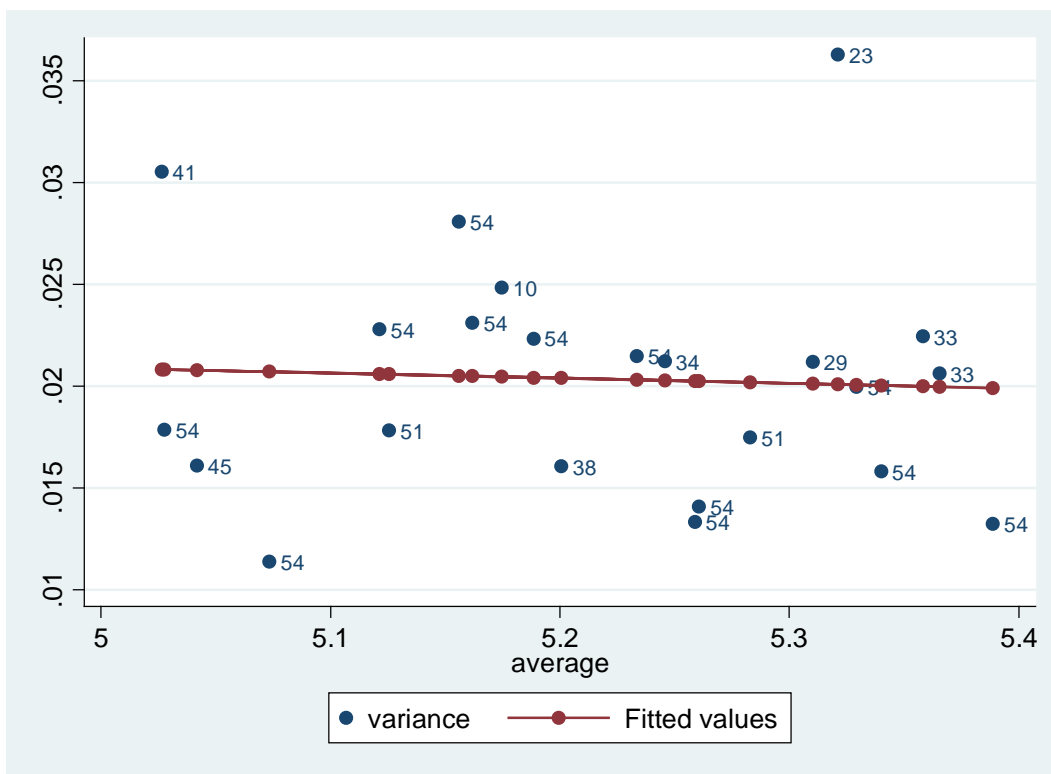
Table 1: Summary of raw bowling scores

player	games	average	sample variance	
player 1	51	169.7843	572.0927288	
player 2	51	198.6863	743.9398081	
player 3	38	182.7632	496.564177	
player 4	41	154.6585	742.3303481	
player 5	34	191.7059	765.0020262	
player 6	10	178.8	897.0665941	
player 7	54	220.3333	615.3961796	
player 8	54	193.963	576.1118454	
player 9	54	189.4444	776.6287682	
player 10	54	193.5	478.9339664	
player 11	54	176.5	758.8961607	
player 12	54	175.8333	846.254863	
player 13	45	155.9778	423.3405011	
player 14	33	216.0606	995.0586347	
player 15	54	160.5926	276.0196795	
player 16	54	169.4259	627.2680539	
player 17	54	181.2037	759.4483244	
player 18	54	153.9259	452.4473581	
player 19	54	208.2963	842.9296056	
player 20	54	210.1852	746.9841076	
player 21	33	214.6364	1006.48895	
player 22	23	208.087	1444.810274	
player 23	29	204.4483	870.3272917	

Table 2: Shapiro Francia Normality Test on the Raw Data

player	W'	V'	P Value	Reject / Dont Reject
Player 1	0.91637	4.404	0.00279	Reject
Player 2	0.91703	4.370	0.00291	Reject
Player 3	0.97441	1.073	0.4466	Dont Reject
Player 4	0.96835	1.407	0.25888	Dont Reject
Player 5	0.97924	0.800	0.66274	Dont Reject
Player 6	0.86434	2.268	0.08115	Dont Reject
Player 7	0.97651	1.294	0.30858	Dont Reject
Player 8	0.92857	3.935	0.00496	Reject
Player 9	0.98555	0.796	0.67344	Dont Reject
Player 10	0.96217	2.084	0.07978	Dont Reject
Player 11	0.95664	2.388	0.04831	Reject
Player 12	0.98831	0.644	0.80833	Dont Reject
Player 13	0.92719	3.478	0.01004	Reject
Player 14	0.95842	1.567	0.20206	Dont Reject
Player 15	0.9825	0.964	0.52841	Dont Reject
Player 16	0.99436	0.311	0.99089	Dont Reject
Player 17	0.96294	2.041	0.08566	Dont Reject
Player 18	0.95042	2.731	0.02803	Reject
Player 19	0.99199	0.441	0.94918	Dont Reject
Player 20	0.94305	3.137	0.01517	Reject
Player 21	0.98768	0.465	0.92728	Dont Reject
Player 22	0.99143	0.248	0.99436	Dont Reject
Player 23	0.97487	0.86	0.61002	Dont Reject
FisherMethod	P= 94.291511	p value = .00003545	df = 46	**Reject**

Table 3: Skewness Kurtosis Normality Test on the Raw Data

player	Pr(Skewness)	Pr(Kurtosis)	adj chi2(2)	P Value	Reject / Dont Reject
Player 1	0.001	0.011	13.79	0.001	Reject
Player 2	0.021	0.620	5.37	0.0681	Dont Reject
Player 3	0.257	0.489	1.88	0.390	Dont Reject
Player 4	0.204	0.853	1.75	0.4168	Dont Reject
Player 5	0.795	0.092	3.16	0.2063	Dont Reject
Player 6	0.052	0.100	5.88	0.053	Dont Reject
Player 7	0.526	0.023	5.39	0.0674	Dont Reject
Player 8	0.002	0.028	11.41	0.0033	Reject
Player 9	0.273	0.662	1.45	0.4838	Dont Reject
Player 10	0.822	0.088	3.12	0.2106	Dont Reject
Player 11	0.033	0.477	4.97	0.0831	Dont Reject
Player 12	0.555	0.515	0.80	0.6712	Dont Reject
Player 13	0.014	0.401	6.23	0.0443	Reject
Player 14	0.233	0.269	2.88	0.2373	Dont Reject
Player 15	0.234	0.777	1.57	0.4571	Dont Reject
Player 16	0.930	0.593	0.29	0.8637	Dont Reject
Player 17	0.124	0.212	4.09	0.1294	Dont Reject
Player 18	0.009	0.081	8.47	0.0145	Reject
Player 19	0.731	0.675	0.29	0.8631	Dont Reject
Player 20	0.012	0.256	6.87	0.0323	Reject
Player 21	0.844	0.312	1.12	0.5698	Dont Reject
Player 22	0.82	0.721	0.18	0.9143	Dont Reject
Player 23	0.812	0.068	3.71	0.1563	Dont Reject
Fisher Method	P=94.718785	p value = .00003152	Reject		

Table 4: Summary Statistics of Transformed Bowling Scores

player	games	average	sample variance
player 1	51	5.125511	0.017810504
player 2	51	5.282947	0.017482155
player 3	38	5.200579	0.016066931
player 4	41	5.02629	0.030541931
player 5	34	5.245745	0.02120932
player 6	10	5.174655	0.024854248
player 7	54	5.388748	0.013224701
player 8	54	5.26058	0.014082023
player 9	54	5.233548	0.02149115
player 10	54	5.258857	0.013321569
player 11	54	5.161812	0.023119233
player 12	54	5.155914	0.028083056
player 13	45	5.041652	0.016096885
player 14	33	5.365455	0.020640351
player 15	54	5.073407	0.011378254
player 16	54	5.121414	0.022802661
player 17	54	5.188546	0.022342387
player 18	54	5.027542	0.017845674
player 19	54	5.32928	0.019949218
player 20	54	5.340073	0.015807329
player 21	33	5.358176	0.022453914
player 22	23	5.321139	0.036293489
player 23	29	5.310156	0.021186666

Table 5: Shapiro-Francia Test on Transformed Bowling Scores

Player	W'	V'	Z	p value	Reject / Dont Reject
Player 1	0.96077	2.066	1.385	0.08308	Dont Reject
Player 2	0.93931	3.196	2.191	0.01423	Reject
Player3	0.95587	1.851	1.152	0.12471	Dont Reject
Player 4	0.98233	0.785	-0.464	0.6786	Dont Reject
Player 5	0.9782	0.84	-0.328	0.6284	Dont Reject
Player 6	0.91056	1.495	0.668	0.25221	Dont Reject
Player 7	0.96957	1.676	0.996	0.15972	Dont Reject
Player 8	0.96354	2.009	1.338	0.09052	Dont Reject
Player 9	0.99301	0.385	-1.917	0.97235	Dont Reject
Player 10	0.9459	2.98	2.071	0.01916	Reject
Player 11	0.9811	1.041	0.078	0.46881	Dont Reject
Player 12	0.98848	0.634	-0.901	0.81618	Dont Reject
Player 13	0.95738	2.036	1.344	0.08941	Dont Reject
Player 14	0.97296	1.019	0.036	0.48575	Dont Reject
Player 15	0.96568	1.891	1.224	0.11056	Dont Reject
Player 16	0.98706	0.713	-0.669	0.74822	Dont Reject
Player 17	0.97846	1.187	0.333	0.36949	Dont Reject
Player 18	0.98105	1.044	0.083	0.46677	Dont Reject
Player 19	0.98905	0.603	-1.002	0.84175	Dont Reject
Player 20	0.9704	1.631	0.943	0.17281	Dont Reject
Player 21	0.98524	0.556	-1.109	0.86639	Dont Reject
Player 22	0.9743	0.743	-0.537	0.70443	Dont Reject
Player 23	0.97535	0.844	-0.315	0.62366	Dont Reject
Fishers Method	P = 61.024478	p value = .06807303	df=46	**Dont Reject**	

Table 6: Skewness Kurtosis Test on Transformed Bowling Scores

player	Pr(Skewness)	Pr(Kurtosis)	adj chi2(2)	P Value	Reject / Dont Reject
Player 1	0.03	0.133	6.4	0.0407	Dont Reject
Player 2	0.071	0.252	4.61	0.0998	Dont Reject
Player 3	0.084	0.9	3.24	0.1976	Dont Reject
Player 4	0.865	0.672	0.21	0.9011	Dont Reject
Player 5	0.745	0.087	3.3	0.1916	Dont Reject
Player 6	0.184	0.228	3.69	0.1578	Dont Reject
Player 7	0.232	0.112	4.12	0.1274	Dont Reject
Player 8	0.048	0.192	5.4	0.067	Dont Reject
Player 9	0.939	0.407	0.72	0.6993	Dont Reject
Player 10	0.095	0.012	7.99	0.0184	Reject
Player 11	0.335	0.871	0.99	0.6095	Dont Reject
Player 12	0.528	0.631	0.65	0.7234	Dont Reject
Player 13	0.085	0.704	3.31	0.1911	Dont Reject
Player 14	0.518	0.147	2.73	0.255	Dont Reject
Player 15	0.056	0.787	3.94	0.1397	Dont Reject
Player 16	0.284	0.936	1.2	0.5492	Dont Reject
Player 17	0.433	0.092	3.64	0.1617	Dont Reject
Player 18	0.159	0.417	2.78	0.2489	Dont Reject
Player 19	0.455	0.763	0.67	0.7166	Dont Reject
Player 20	0.116	0.592	2.9	0.2346	Dont Reject
Player 21	0.605	0.316	1.36	0.5074	Dont Reject
Player 22	0.258	0.879	1.42	0.4909	Dont Reject
Player 23	0.789	0.082	3.41	0.1814	Dont Reject
Fisher Method	P = 68.695722	p value = .01666579	Reject		

Table 7: Probabilities of favourite winning unadjusted scores and adjusted scores under handicap systems 1 and 6

Ability	100	120	140	160	180	200	220
100		.8184	.9527	.9904	.9985	.9998	1
		.5665	.617	.661	.6981	.7332	.7566
		.6178	.7059	.7672	.827	.8680	.8913
120	8184		.7775	.9236	.9812	.9962	.9989
	.5665		.5604	.5963	.6411	.6800	.7035
	.6178		.6036	.6761	.7394	.7944	.8332
140	.9527	.7775		.7430	.8959	.9628	.9882
	.617	.5604		.5461	.5920	.6308	.6579
	.7059	.6036		.5835	.6567	.7239	.7683
160	.9904	.9236	.7430		.7299	.8722	.9514
	.661	.5963	.5461		.5457	.5845	.6207
	.7672	.6761	.5835		.5808	.6462	.7033
180	9985	.9812	.8959	.7299		.7063	.8503
	.6981	.6411	.5920	.5457		.5419	.5797
	.827	.7394	.6567	.5808		.5754	.6346
200	.9998	.9962	.9628	.8722	.7063		.6848
	.7332	.6800	.6308	.5845	.5419		.5348
	.8680	.7944	.7239	.6462	.5754		.5633
220	1	.9989	.9882	.9514	.8503	.6848	
	.7566	.7035	.6579	.6207	.5797	.5348	
	.8913	.8332	.7683	.7033	.6346	.5633	

Table 8: Probabilities of favourite winning unadjusted and New handicap system based on 100% factor weighting

Ability	100	120	140	160	180	200	220
100		.8184 .4958	.9527 .4841	.9904 .4773	.9985 .4772	.9998 .4805	1 .4747
120	.8184 .4958		.7775 .4984	9236 .4838	.9812 .4806	.9962 .4833	.9989 .4775
140	.9527 .4841	.7775 .4984		.7430 .4918	.8959 .4891	.9628 .4908	.9882 .4836
160	.9904 .4773	.9236 .4838	.7430 .4918		.7299 .4934	.8722 .4963	.9514 .4907
180	.9985 .4772	.9812 .4806	.8959 .4891	.7299 .4934		.7063 .4973	.8503 .4949
200	.9998 .4805	.9962 .4833	.9628 .4908	.8722 .4963	.7063 .4973		.6848 .4952
220	1 .4747	.9989 .4775	.9882 .4836	.9514 .4907	.8503 .4949	.6848 .4952	

Table 9: Probabilities of favourite winning unadjusted and New handicap system based on 95% weighting factor "Adjusted score = rawscoreplayer+.95(220-m_i)"

Ability	100	120	140	160	180	200	220
100		.8184 .5155	.9527 .5184	.9904 .5264	.9985 .5330	.9998 .5465	1 .5472
120	.8184 .5155		.7775 .5152	.9236 .5133	.9812 .5196	.9962 .5311	.9989 .5353
140	.9527 .5184	.7775 .5152		.7430 .5039	.8959 .5137	.9628 .5291	.9882 .5268
160	.9904 .5264	.9236 .5133	.7430 .5039		.7299 .5057	.8722 .5187	.9514 .5243
180	.9985 .5330	.9812 .5196	.8959 .5137	.7299 .5057		.7063 .5080	.8503 .5186
200	.9998 .5465	.9962 .5311	.9628 .5291	.8722 .5187	.7063 .5080		.6848 .5050
220	1 .5472	.9989 .5353	.9882 .5268	.9514 .5243	.8503 .5186	.6848 .5050	