Communication in Mathematical Modeling and Applications
http://ntmsci.com/cmma

Testing the consistency of a novel bowling scoring system against the current approach

Abdel-Salam G. Abdel-Salam, Mohammad D.Mollazehi, IbrahimA. El Feky, Yousuf A. Dorra, Irfan J. Ismail and Salman Umer

Department of Mathematics, Statistics, and Physics, College of Arts and Sciences, Qatar University, Qatar

Received: 27 August 2017, Accepted: 21 October 2017
Published online: 22 October 2017.

Abstract: This study introduces a new bowling scoring system and tests its consistency with the current scoring system with respect to preserving player placements. A comprehensive simulation study for different scenarios; two-player, three-player, and four-player games performed. The simulations empirically quantify the likelihood of experiencing concordant results between the two (current and new) systems. The simulation study revealed that the percentage of times that the current and new scoring systems yield the same placements at least 85% when two players compete, at least 66% of the time when three players compete, and at least 43% of the time when four players compete regardless of the ability levels of the players. A comparison study using real bowling data-sets have been done and showed the consistence for the new scoring system with the current one. However, the new scoring system is easy to calculate, understand and to be implemented.

Keywords: Scoring System, bowling, conditional modeling, categorical data analysis, cumulative logit model.

1 Introduction

Bowling is a fun activity that allows people to gather and enjoy knocking down as many pins as possible. Bowling can be played in a competitive or leisure fashion, and is often played with two or more players. The object is to bowling is to ‘bowl’ as many pins as possible to achieve the highest score amongst the other players. Although the activity is considered “fun”, understanding its scoring system can be challenging to amateur bowlers. Rogers (2014) identified that bowling has an ‘image problem’ due to its scoring system. When a player fails to bowl all ten pins down in a frame, scoring is trivial since the score involves simple addition. For instance, say that a player bowls a total of eight pins in the frame and seven in the second frame. The total score up to that point would be $8+7=15$ points. However, the scoring rubric become challenging when strikes or spares are involved (i.e. two ways of successfully bowling all ten pins in a player’s turn). For instance, bowling all ten pins in the first frame followed by only six pins in the second frame does not equate to 16 points. Rather, the total is 22 points. This is because a bonus is awarded whenever a player bowls all ten pins in one frame. Amateur bowlers often forget this incentive when playing. The scoring system becomes even more cumbersome when a player achieves consecutive strikes, consecutive spares, or a combination of both.

Several sources suggest that the history of bowling dates back to as early as 3200 and 5200 BC (Bowling, n.d.; Pretsell, 1908), but records of bowling competitions began mostly after the development of bowling associations and bowling magazines in the 1800s (Pluckahn, 1988). Regardless of the reason, the object is to throw a bowling ball down a lane towards ten bowling pins in an attempt to knock them down. The more pins that are knocked down, the higher the bowling score. Bowling can be played individually or competitively using two or more players. When played competitively, the winner is determined based on the highest score. Bowling requires some level of skill and precision. According to Mullen (2004), a bowler’s stance, approach and timing in rolling, the footwork, arm swing, and delivery

* Corresponding author e-mail: Abdo@qu.edu.qa

contribute to achieving a high score.

According to Ryan (2016), there are several varieties of bowling including ten-pin, no-tap, Candlepin, Duckpin, and five-pin bowling. This study focuses on ten-pin bowling since it is considered the most popular version with rules structured by bowling associations in the United States and Canada (Harmon, 1985). With ten-pin bowling, the bowling ball contains three to four finger holes and weighs between 2.72 to 7.26 kilograms. The ball is rolled towards standing pins with a height of 38.1 cm and a weight between 1.42 and 1.64 kilograms. The pins are placed at the far end of the alley in the form of a triangle. For more information on the dimensions of equipment used in ten-pin bowling, refer to Bowling (n.d.) or Mullen (2004).

A bowling game consists of ten frames. Within each of the first nine frames, two rolling scores and one bowling score is recorded. In the tenth frame, two or three rolling scores and the final bowling score are recorded. Details on how to record the rolling and bowling scores follow. It is important to the rolling scores represent how well the bowler knocked down the ten pins in each of the frames, but the bowling scores are recorded as cumulative scores from the first frame to the tenth frame.

From the aforementioned, a bowler's objective is to knock down all ten pins per frame within one or two rolls. When he or she completes his or her frame, there are three possible outcomes. One outcome is that the bowler successfully knocks all ten pins down on the first roll. This is called a *strike frame* and the notation used as the score is an "X". The score for the strike is recorded as the first rolling score and no score is entered for the second rolling score. The second outcome is called a *spare frame*, that is the bowler successfully knocks all ten pins down in two rolls. For example, he or she knocked down eight pins in the first roll followed by the remaining two pins in the second roll. Here, the rolling scores are recoded as an "8" and "/" respectively. The symbol "/" indicates that the spare was successful, meaning that the remaining two pins were knocked down on the second roll. The final outcome is an *open frame*. This is when a bowler failed to knock down all ten pins in two rolls, more specifically at least one pin remained standing. For example, a bowler could knock down seven pins in the first roll and two of the three remaining pins in the second roll. The rolling scores would be recorded as "7" and "2' respectively. Note that for strike and open frame outcomes, the word "frame" is typically removed from this term.

The tenth frame is somewhat similar to the behavior of the ninth frame. A bowler is given two rolls. If the two rolls result in an "open frame", the frame ends. However, if a bowler successfully rolls a strike within the first roll or a spare within two rolls, an additional roll is awarded.

There has been recent events involving the creating of new bowling scoring systems. Rogers (2014) addressed that a NSS was used at the 2014 World Bowling Tour with the purpose of increasing game suspense and the speed of play. However, this scoring system altered the bowling game play since it advanced bowlers based on a "frame-by-frame showdown" in which one roll is made per frame. This is distinct with the scoring system in this study where bowlers continue to use two rolls per frame. Rogers (2014) also emphasizes that the NSS, developed by the international organization *World Bowling*, would attract the attention of the International Olympic Committee in order to make bowling an Olympic sport as soon as 2024.

Statistical research studies on bowling scoring systems are nearly nonexistent. Regrettably, the most significant and relevant research on scoring systems investigated the effects of the handicapped bowling on scoring (Keogh and O'Neill, 2011). These researchers examined the impact of alternative handicapped scoring after applying a log transformation on bowling data captured in Ireland. Keogh and O'Neill (2011) found that the handicap system system allows lower-ability players (i.e. those with a handicap) to compete fairly across all match-ups against those without a handicap. Other statistical research investigated winning-streak patterns in bowlers' performances over several tournaments called the "hot hand phenomenon" (Martin (2006); Yaari and David (2012)). These studies address the consistency of performances at the bowler level rather than the consistency of scoring systems. Also, other research in bowling studies the distribution of final bowling scores under the CSS. Neal (2003) created an algorithm using Mathematica to simulate and understand the average random bowling score. He found that, with random rolling of the bowl and its outcomes, the average score per frame was approximately 9.14 points (out of 30). Cooper and Kennedy (1990), on the other hand, used the Pascal programming language to compute the number of ways of achieving a particular final score in bowling. They found, for example, that there are 50, 613, 244, 155, 051, 856 ways to create a score of 100 points using the current system, but only a little over 1,000 ways of achieving a score of 267 points. Also, Hohn (2009) introduced a method to describe the distribution of scores per frame "for N frames and M pins per frame" and across various ability levels.

In this study, the researchers created an innovative bowling scoring system that is believed to be easier to understand than the current scoring system. Rather than determine whether the system is easier to understand, the purpose of this study is to quantify the consistency between the two scoring systems. This mean establishing how often the bowler placements are similar under this system compared to the current scoring system used in bowling. There are three main research questions:

(1) What percentage of times do the two scoring systems give identical placements?
(2) How is the percentage affected by the number of players?
(3) Can this behavior be modeled?

The remainder of this paper is organized as follows. We first review the current scoring system for bowling. We then describe the use of mixed model penalized spline (p-spline) regression as a NP method to estimate the nonlinear mixed effects profiles. Next, we introduce the nonlinear mixed robust profile monitoring $(NMRPM)$ method. In addition, we present diagnostic tools to determine outlying profile(s) for our new methods. A real application study using the dose-response dataset discussed earlier is utilized to compare the new proposed methods to the existence parametric nonlinear mixed models. We conclude by giving conclusions and future research ideas.

2 Current scoring system

Under the current scoring system (CSS), the number of points awarded per frame is based on its outcome (i.e. strike frame, spare frame, or open frame). The calculation of the number of points for an open frame is trivial since it is the sum of the two rolling scores. For a spare frame, the number of points awarded is delayed until the outcome of the first roll in the next frame. After which, the number of points awarded is the sum of 10 points and the number of points of that first roll. For a strike frame, the number of points awarded is delayed until the outcome of the next two rolls that may occur in the next frame or the next two frames. After which, the number of points awarded is the sum of 10 points and the number of points of those two rolls.

Frame 1	Frame 2	Frame 3	Frame 4	Frame 5	Frame 6	Frame 7	Frame 8	Frame 9	Frame 10
X	4 /	5 3	8 /	X	X	X	9 -	5 /	7 / 8
20	35	43	63	93	122	141	150	167	185

Fig. 1: CSS example

Figure 1 illustrates the results of a bowling game for a player. The first frame shows an example of a strike frame. The score is withheld until the next two rolls, for which both of them took place in Frame 2. The total of the two immediately-rolled frames add up to 10 points. So, the score for Frame 1 is calculated by adding 10 points for the strike and 10 more points for the spare (i.e. 10 points + 10 points = 20 points). Frame 2 is an example of a spare frame. As a result, the score is withheld again until the outcome of the first roll (occurring in Frame 3). Since the next roll shows that five pins were knocked down, the total number of points awarded for Frame 2 is 10 points (for the spare) + 5 points = 15 points. Frame 3 is an example of an open frame for which a total of eight pins were knocked down. This, therefore, becomes the total number of points awarded for this frame. The outcome for Frame 4 is similar to that of Frame 2, with the exception of ten additional points being awarded for the strike in the immediate roll. Thus, twenty points are awarded for Frame 4. Frame 5 is similar to that of Frame 1 where twenty additional points were added to the ten points for the strike (i.e. 10 points for the strike in Frame 6 and 10 more points for the strike in Frame 7). Thus, the bowler receives a total of 30 points for Frame 5. Frame 6 results in a total of 29 points (10 points for the strike in Frame 6, 10 points for the strike in Frame 7 and 9 points from the first roll of Frame 8). One can show that 19 points is awarded for Frame 7, 9 points is awarded for Frame 8, and 17 points for Frame 9. For Frame 10, the bowler receives a bonus roll for successfully rolling a spare. The additional roll resulted in the bowler knocking down eight of the ten pins. The total number of points earned from Frame 10 is equal to 18 (i.e. 10 points for the spare + 8 points for the third roll).

Bowling scores are recorded cumulatively from the first to the tenth frame, that is the total number of points are added as the bowlers completes each frame. The tenth frame contains the final bowling score. The minimum possible final score is 0 points. This can happen when a bowlers was unsuccessful in knocking down any pins for all ten frames. The maximum final score is 300 (called a “perfect game”), which can occur if a bowler successfully rolls twelve consecutive strikes. In the follwoing Section, the new scoring system will be introduced.

3 New scoring system

The new scoring system (NSS) involves rewarding scores that are in powers of ten for open, spare, and strike frames. Similar to the current system, it rewards players with extra points for consecutive spares and strikes. There are two distinct advantages of the NSS. First, it is designed to make calculating one’s score easier since adding values of 0, 1, 10, 100, or 1,000 is simpler. Second, the NSS is more ‘real-time’ than its current counterpart, namely that one’s score can be calculated immediately after completing his or her frame.

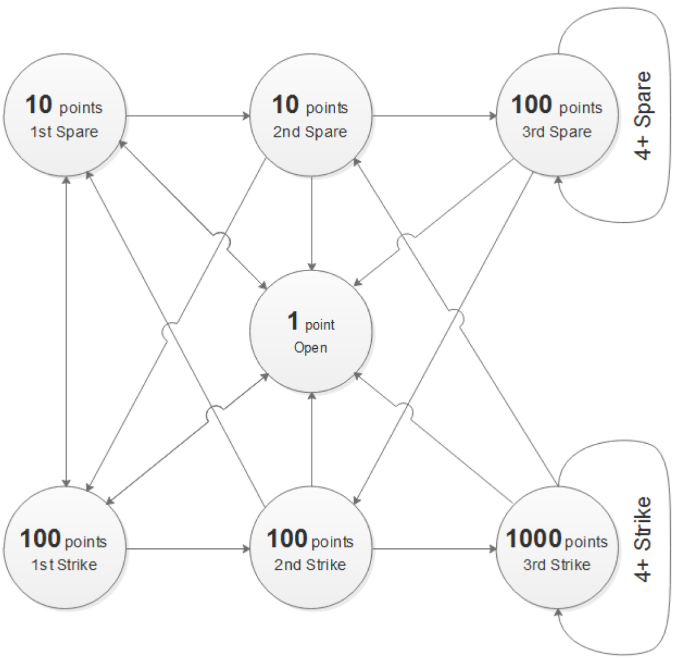

Fig. 2: NSS diagram.

Figure 2 outlines how the NSS works. Unlike the CSS which assigns points based on the actual number of pins bowled, the NSS assigns points based on whether all of the pins are knocked down or not. As discussed in Section 2, a frame is categorized as either an *open frame*, a *spare frame*, or a *strike frame*. For an open frame, a score of 0 points is given to a bowler when no pins are knocked down. Note that this score is not mentioned in Figure 2 since it is implied that no points should be awarded when no pins are knocked down. However, a score of only 1 point is given if **at least** one pin is knocked down. More points are awarded when a bowler successfully knocks all ten pins in a frame using one or two rolls. For the first spare frame, 10 points are awarded. For the first strike frame, 100 points are awarded. Similar to the CSS, more points are awarded when a bowler performs three or more consecutive spare frames or three or more consecutive strike frames. Each spare after the second consecutive spare results in 100 points. Similarly, 1,000 points are awarded for each consecutive strike after the second strike. It is important to emphasize that bonus points are not awarded when a bowler interchanges between spares and strikes between frames. For example, if a bowler rolls two consecutive strikes followed by a spare and then another strike, the bowler did not receive 1,000 points on the third strike since the strikes were not consecutive (i.e. “breaking the chain”).

Note that Figure 2 shows links between spare and strike frames. For example, four consecutive spares followed by a strike would result in awarding 10 points (for the first spare), 10 points (for the second spare), 100 points (for the third spare), another 100 points (for the fourth spare), and 100 points (for the “second-level” strike) respectively. This means that a bowler can receive 1,000 points for the next strike rather than needing three consecutive strikes to achieve the same bonus. This is referred to as the “fall-back” rule. The fall-back rule allows bowlers some level of an advantage towards getting to the third-level bonus points without having from the first level. The fall-back rule does not apply when a bowler rolls an open frame. When this happens, a bowler must start over in developing consecutive spares or consecutive strikes to get to the 100 or 1,000 bonus points respectively. Finally, arrows appear in Figure 2 to show all of the possible outcomes between frames. For example, the two arrows connecting the first strike frame score to the first-level open frame score implies that a bowler can roll a strike frame and then a open frame, and vice versa. Note, however, that there is a one-way arrow connecting the third-level strike frame score to the open-frame score. This means that a bowler can transition from a third-strike frame to an open frame between two consecutive frames, but cannot transition from an open frame immediately to a third-level strike frame between two consecutive frames. In this case, he or she must work up the chain again. The same rules imply for the spare frame case.

The same scoring rubric applies to score the tenth frame. If an open frame occurs and the bowler does not use the bonus roll, then only 1 point is awarded for the tenth frame. If a bowler rolls a spare or a strike and uses the bonus roll, then the bowler receives points for the spare and points for whatever occurs from the bonus roll. Note that a bowler can carry advantages from previous frames into the tenth frame. Thus, he or she can reach the third-level bonus points from achieving three or more consecutive spares or strikes.

Similar to the CSS, bowling scores are recorded cumulatively from the first to the tenth frame. The total number of points are added as a bowler completes each frame. The tenth frame contains the final bowling score. The minimum possible final score is 0 points, which occurs when a bowler unsuccessful knocks down any pins for all ten frames. The maximum final score is 10,200, which can occur if a bowler successfully rolls twelve consecutive strikes. It is important to note that the rules and structure of bowling do not change due to the NSS.

Frame 1		Frame 2		Frame 3		Frame 4		Frame 5		Frame 6		Frame 7		Frame 8		Frame 9		Frame 10		
X		4	/	5	3	8	/	X		X		X		9	-	5	/	7	/	8
100		**110**		**111**		**121**		**221**		**321**		**1,321**		**1,322**		**1,332**		**1,343**		

Fig. 3: NSS example

Figure 3 is an example of how to calculate the final score under the NSS using the same frame outcomes from Figure 1. The outcome from Frame 1 was a first-level strike, and so 100 points are awarded. Frame 2 shows a first-level spare, and so 10 points are awarded. Because of the open frame in Frame 3, only 1 point is awarded. Frame 4 is an example of a first-level spare again since it was not a part of consecutive spares. So, only 10 points are awarded. Frames 5, 6, and 7 show three consecutive strikes. So, 100 points are awarded for the first-level strike in Frame 5, 100 points for the second-level strike in Frame 6, and 1,000 points for the third-level strike in Frame 7. Frame 8 is an example of an open frame, and so only 1 point is awarded. Frame 9 is a first-level spare resulting in 10 points. For Frame 10, the bowler rolled a second-level spare followed by an open frame. This means that 11 points are awarded (i.e. 10 points for the second-level spare and another 1 point for the open frame due to the bonus roll). The final score for this bowler is 1,343 points, which is in contrast to the 185-point final score under the CSS.

3.1 The two scoring systems and placements

The scoring systems described in Sections 2 and **??** clearly demonstrate two distinct methods for determining the placements of bowlers in a bowling game. While the CSS and NSS yield higher scores for better bowling performance, it is important to emphasize that the two methods are not perfectly consistent. Without loss of generality, the proceeding examples that show how they can be consistent or inconsistent using three-player bowling games.

Bowler	**Final Score**		**Placements**	
	Current	New	Current	New
A	171	542	1	1
B	130	226	2	2
C	81	20	3	3

Table 1: Consistent placement results.

Table 1 shows the placements of the three bowlers under the CSS and NSS. Placements are determined based on the final bowling score. It is left to the reader to show that the final scores under the NSS for the three bowlers are 542 (Bowler A), 226 (Bowler B), and 20 (Bowler C) respectively. These final scores mean that Bowler A has first place, Bowler B has send place, and Bowler C has third place. Therefore, the order of the placements for this game is (1,2,3) for both scoring systems.

Bowler	**Final Score**		**Placements**	
	Current	New	Current	New
A	120	137	1	2
B	117	1,207	2	1
C	109	46	3	3

Table 2: Inconsistent placement results.

However, one can see that the final scores under the NSS for the three bowlers are 137 (Bowler A), 1,207 (Bowler B), and 46 (Bowler C). These final scores mean that Bowler B has first place, Bowler A has send place, and Bowler C has third place. Table 2 shows the placements of the three bowlers under both scoring systems. Based on the final scores, the order to the placements for this game are (1,2,3) for the CSS and (2,1,3) under the NSS. Note that the placements are now inconsistent. There are two plausible reasons for this phenomenon. One reason is the closeness of the bowling scores. The differences in placements between the two games are different. When scores are fairly close, the placements could differ.

A second plausible reason is the performance within the frames. Bowlers A and C managing multiple spares but never achieved a third consecutive spare to receive the largest bonus. Bowler B, however, successfully rolled three consecutive strikes to receive a 1,000-point bonus and never achieved any spares throughout the game.

4 Simulation study

The simulation study consisted of creating six different datasets that differ by the number of bowlers and whether the bowlers have equal or unequal ALs. Each simulated dataset contained 10,000 games. The study contained three independent variables. The first independent variable was the number of bowlers (NB). The levels for this variable were two players (Bowlers A and B), three players (Bowlers A, B, and C), and four players (Bowlers A, B, C, and D). The second independent variable was the ability level (AL) of the bowlers. We considered the case of bowlers having equal ($\theta_A = \theta_B = \theta_C = \theta_D$) and unequal ($\theta_A < \theta_B < \theta_C < \theta_D$) bowling abilities where θ is the AL. The third independent variable was the placings of the bowlers under the CSS. Possible placings were 1 (first place for having the highest score), 1.5 (tie between the first and second places), 2 (second place), 2.5 (tie between second and third places), 3 (third place), 3.5 (tie between third and fourth places), and 4 (fourth place for having the lowest score). All simulated data were created using the SAS 9.4 statistical software.

The dependent variable was the placings of the bowlers under the NSS. Similar to that of the CSS, possible placings were 1 (first place for having the highest score), 1.5 (tie between the first and second places), 2 (second place), 2.5 (tie between second and third places), 3 (third place), 3.5 (tie between third and fourth places), and 4 (fourth place for having the lowest score).

For each game, the outcome of the rolls were generated for ten frames similar to that of real bowling games. Table 3 contains the parameters used to create the rolling scores for the four-player games. For example, consider the case where the bowlers have unequal ALs. For Bowler A, the B(10,0.65) for Roll 1 (R_1) means that the score for the first roll is generated using a binomial distribution with n = 10 and probability p = 0.65 of knocking down each of the ten pins. The B(10-R_1,0.70) means that, of the remaining pins standing, generate the score for Roll 2 using a binomial distribution of n = 10 - R_1 and a probability p = 0.70 of knocking down each of those pins. The probability for the second roll was set to be higher than the first roll since it is believed that a bowler will be more focused in knocking down the remaining pins to achieve the spare if some pins remained after the first roll. Note that the probabilities increase from Bowler A to Bowler D. This is obvious since it was stated earlier that $\theta_A < \theta_B < \theta_C < \theta_D$ for the unequal AL case.

For the equal ability case, it is obvious that all of the probabilities must be the same value. The probabilities of p = 0.725 for the first roll and p = 0.775 for the second roll were considered since they are the averages of the four probabilities from the unequal ability case for Rolls 1 and 2 respectively.

The final bowling scores under the CSS and NSS were calculated using the scores from the simulated rolls. Then, the placings were calculated for each bowler per game under the CSS and NSS.

Bowlers	**Unequal Ability**		**Equal Ability**	
	Roll 1 (R_1)	Roll 2	Roll 1 (R_1)	Roll 2
A	$B(10,0.65)$	$B(10-R_1,0.70)$	$B(10,0.725)$	$B(10-R_1,0.775)$
B	$B(10,0.70)$	$B(10-R_1,0.75)$	$B(10,0.725)$	$B(10-R_1,0.775)$
C	$B(10,0.75)$	$B(10-R_1,0.80)$	$B(10,0.725)$	$B(10-R_1,0.775)$
D	$B(10,0.80)$	$B(10-R_1,0.85)$	$B(10,0.725)$	$B(10-R_1,0.775)$

Table 3: Simulation parameters for four-player games.

It is the researchers' interests to investigate whether the behavior of the placements can be explained using a regression model. Because the placements of the NSS is an ordinal categorical variable with more than two possible values, one appropriate model to use is the cumulative logit model. For this study, the cumulative logit regression model has the form

$$logit[P(Y \leq k)] = \alpha_k + \beta x, \tag{1}$$

where

$$logit[P(Y \leq k)] = \log\left[\frac{P(Y \leq k)}{1 - P(Y \leq k)}\right], \quad k = 1, 1.5, 2, 2.5, \ldots, K - 0.5. \tag{2}$$

and K is the highest possible placement under the NSS. Here, K = 2 when considering a two-player game, K = 3 for three players, and K = 4 for four players. The cumulative logit model is preferred over the linear regression model because the research questions imply modeling the likelihood of obtaining a particular placement and not modeling the actual placement.

There were two conditions to consider in order to adequately apply the cumulative logit model. To make the regression model easier to understand, the models used those games in which there were no ties in the placements for the CSS only. This prevented the researchers from violating the independence of observations rule. Another requirement to control the independence of observations was to create conditional cumulative logit models for each placement under the CSS. This helped separate the observations into distinct datasets that are independent. By doing so, AL was the only independent variable used in the model.

Four main-effects cumulative logit models were produced for the four-player case, three for the three-player case, and only one for the two-player case. An intercept-only model was also produced with each cumulative logit model. For example, consider the four-player models. Model 1b was the main-effects model that predicted the likelihood of having the $k^t h$ placement given the AL and placing first under the CSS. Model 1a was its intercept-only model. Model 2a and 2b were the intercept-only and main-effects models for predicting the likelihood of having the $k^t h$ placement given the AL and placing second under the CSS respectively. Models 3a and 3b would be conditional on placing third under the CSS, while Models 4a and 4b were for placing fourth under the current system.

Several statistics were presented with the model estimates for the main-effects models. The deviance (-2 Log L) and Akaike Information Criterion (AIC) statistics help compare the main-effects model to its respective intercept-only model. The are popularly-used statistics to help quantify the amount of error explained after considering AL in the regression model. The correct classification rate is a measure of predictive power by determine what percentage of observations were correctly predicted by the model. The Score's Test statistic helped determine whether the proportional odds assumption was met, meaning whether the slope remained constant across all levels.

4.1 Simulation's results

NB	**AL**	**Speraman's Correlation**
2	Equal	0.8272
	Unequal	0.8564
3	Equal	0.8400
	Unequal	0.8901
4	Equal	0.8346
	Unequal	0.9072

Table 4: Correlations between CSS & NSS final scores.

Table 4 shows the Spearman correlations between the final scores from the CSS and NSS. Spearman correlation values ranged between 0.8272 and 0.9072. Note that the highest correlations occurred for the unequal AL cases. All combinations for the number of players and ALs experienced correlations well above the 0.7 threshold. This shows evidence of a strong linear relationship between the two scoring systems.

There were a few recognizable patters from these results. First, the percentage of concordant results decreased as the number of players increased. Second, the percentages for the unequal AL cases were higher than those of the equal ability cases. The percentage of partially-concordant placements increased as the number of players increased. This is trivial since the number of players considered increased the number of possible placements.

NSS Placements	2 Players		3 Players		4 Players	
	Equal	Unequal	Equal	Unequal	Equal	Unequal
1	82.09%	84.46%	73.78%	80.10%	67.74%	76.73%
1.5	2.94%	2.50%	3.79%	2.50%	3.12%	2.57%
2	14.97%	13.04%	17.85%	15.26%	20.95%	17.74%
2.5	-	-	1.24%	0.50%	1.41%	0.57%
3	-	-	3.34%	1.64%	5.46%	2.15%
3.5	-	-	-	-	0.20%	0.03%
4	-	-	-	-	1.11%	0.20%

Table 5: Percentage of placements for NSS given 1st place for CSS.

Table 5 contains conditional percentages for each possible placement under the NSS given the bowler placed first under the CSS. The conditional percentages are provided for all combinations of the number of players and ALs. For example, consider the two-player games with equal ALs of the bowlers. Of the games in which there was a first-place bowler, 82.09% of those bowlers also placed first under the NSS, another 2.94% of bowlers were tied for first and second places, and the remaining 14.97% of bowlers placed second. There were some similar patterns with this table. First, it is no surprise that the highest percentages for this table occurred with the first placement. This, therefore, suggested that one is more likely to maintain first place under both scoring systems. Second, the highest percentages for maintaining first place occurred when the bowlers had unequal versus equal ALs. Similarly, conditional percentages for each possible placement under the NSS given the bowler placed second under the CSS.

4.2 Simulation study- cumulative logit regression models

There were three sets of regression models: one set analyzing the results of the two-player games, a second set analyzing the three-player games, and a final set for the four-player games. Each set contains an intercept-only model containing no variables in the model, a main-effects model containing the AL and the current scoring system placement (CSSP), and a reduced model which eliminated any variables that were not significant. Model error statistics include the deviance statistic (-2 Log L) and the Akaike Information Criterion (AIC). Each set will be described in the following subsections.

4.2.1 Two-player model results

Estimates	Intercept-Only $\hat{\beta}$	Intercept-Only $SE_{\hat{\beta}}$	Main-Effects $\hat{\beta}$	Main-Effects $SE_{\hat{\beta}}$	Reduced $\hat{\beta}$	Reduced $SE_{\hat{\beta}}$
Intercept ($\hat{\alpha}$)						
$\hat{\alpha}_1$	-0.06		1.60		1.60	
$\hat{\alpha}_{1.5}$	0.06		1.82		1.82	
Ability Level (AL)						
$AL_{unequal}$			< 0.001	0.03		
CSS Placement (CSSP)						
$CSSP_{1.5}$			-1.71***	0.08	-1.71***	0.03
$CSSP_2$			-3.43***	0.03	-3.43***	0.03
-2 Log L	64,264.49		42,689.29		42,689.29	
AIC	64,268.49		42,699.29		42,697.29	
CCR			82.63%		82.63%	
Score's Test			27.45 (df = 3)***		24.52 (df = 2)***	

*$p < 5\%$,**$p < 1\%$, ***$p < 0.1\%$

Table 6: Two-player cumulative logit model estimates (simulation).

Table 6 contains the model estimates of the cumulative logit model for the two-player games. For the intercept-only model, the deviance and AIC statistics were 64,264.49 and 64,268.49 respectively. These values represent a base for understanding the amount of unexplained error when no explanatory variables are used to predict NSS placement. The deviance and AIC statistics for the main-effects models were much lower than those from the intercept-only model (-2 Log L = 42,689.29 and AIC = 42,699.29). Based on the slope estimate, AL was not a significant factor in predicting NSS placements ($\hat{\beta}_{AL_{unequal}}$ = $<$ 0.001, $p_v > 5\%$). Because the value of this slope was close to zero, this suggested that two-player games containing equal or unequal ability levels have nearly no effect on NSS placement outcomes and their predicted probabilities. However, CSSP was strongly significant in predicting NSS placement. The estimates of $\hat{\beta}_{CSSP_{1.5}}$ = -1.71 and $\hat{\beta}_{CSSP_2}$ = -3.43 suggested that predicted probabilities for first place under the NSS decreased as the CSSP increased. This is strong evidence suggesting that the CSS and the NSS are highly consistent. The CCR for the main-effects model was 82.63%, showing that 82.63% of NSS placements from players in the simulated two-player games were correctly predicted by the regression model. The main-effects model failed to satisfy the proportional odds assumption (ST = 27.45, df = 3, $p_v < 0.1\%$), resulting in the possibility of negative probabilities. However, it was checked and verified that no negative probabilities exist. The reduced model contained only CSSP as the explanatory variable since AL was not significant. Very little change in results occurred after removing AL. Deviance and AIC statistics were similar (-2 Log L = 42,689.29 and AIC = 42,697.29). CSSP remained significant and contained the same slope estimate values, meaning that that predicted probabilities for first place under the NSS decreased as the CSSP increased. CCR and the Score's Test statistic also remained nearly unchanged compared to that of the main-effects model.

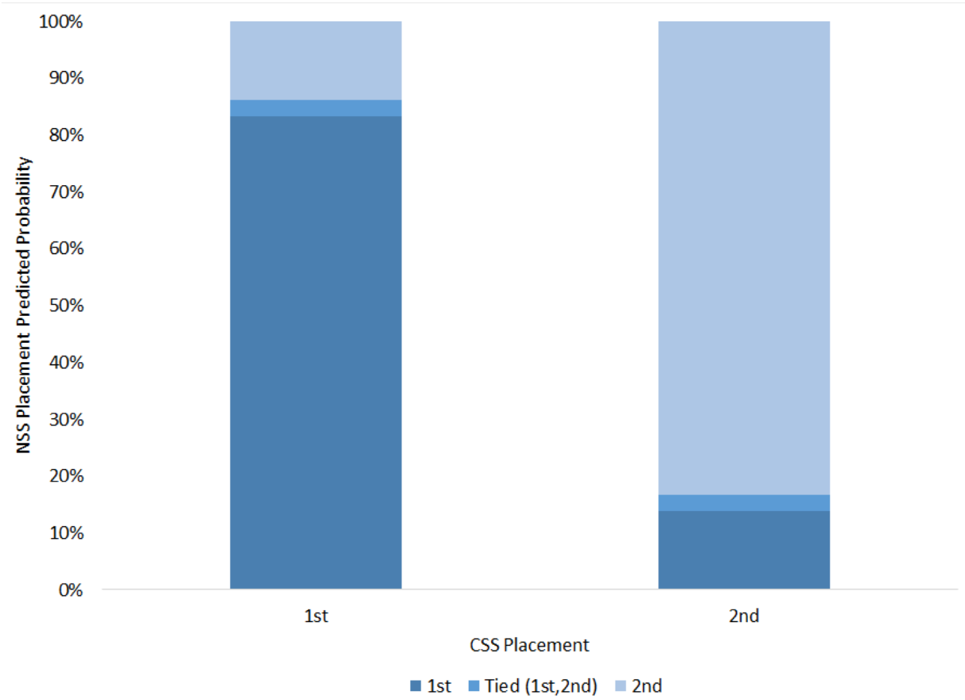

Fig. 4: Two-player predicted probabilities from cumulative logit model.

Figure 4 contains a bar graph showing the predicted probabilities of each NSS placement given a player's CSS placement for the two-player games. The base bar represents the predicted probability of coming in first place under the new scoring system, the middle bar is the predicted probability for a tie between first and second place, and the top bar is for the predicted probability of placing second. For example, the predicted probability of placing first under the NSS was 83.26%, 2.82% for a tie between first and second place, and 13.92% for placing second. it can be clearly seen from Figure 4 that first-place players from the CSS are more likely to come in first place again under the NSS, while second-place players are more likely to remain in second place under the NSS. This was explained from the negative slopes for CSS from the main-effects and reduced models of Table 6.

4.2.2 Three-player model results

Table 7 contains the model estimates of the cumulative logit model for the three-player games. The interpretation of all statistics in this table is similar to that of Table 6. The deviance and AIC statistics were 157,743.00 and 157,751.00. These values decrease sharply when including AL and CSSP to create the main-effects model (-2 Log L = 108,213.72 and AIC = 108,231.72). According to the main-effects model, AL was not significant in predicting NSS placement ($\hat{\beta}_{AL_{unequal}} = <$ 0.003, $p_v > 5\%$). Therefore, this suggested that it is irrelevant whether the three players are equally good or not. The CSSP slopes were all negative, meaning that the chances of coming in first place under the NSS continues to decrease compared bowlers that come in first place under the CSS. This pattern is the same for the reduced model since the intercepts and the slope estimates did not change compared to the main-effects model, and the AIC changed very little after removing AL from the model. The CCR was 71.40%, which means that the regression model correctly predicted NSS placements for 71.40% of the bowlers in the data. This CCR is much smaller than that from the two-player model. One possible reason for this is because with more players playing comes more possible placements, and with more possible placements comes more possible ways to introduce variability in the placements in the data. The simulation results (not presented) showed the same behavior as 2 and 3 players games.

Estimates	Intercept-Only $\hat{\beta}$	Intercept-Only $SE_{\hat{\beta}}$	Main-Effects $\hat{\beta}$	Main-Effects $SE_{\hat{\beta}}$	Reduced $\hat{\beta}$	Reduced $SE_{\hat{\beta}}$
Intercept ($\hat{\alpha}$)						
$\hat{\alpha}_1$	-0.75		1.19		1.19	
$\hat{\alpha}_{1.5}$	-0.64		1.39		1.39	
$\hat{\alpha}_2$	0.60		3.86		3.87	
$\hat{\alpha}_{2.5}$	0.79		4.27		4.27	
Ability Level (AL)						
$AL_{unequal}$			0.003	0.02		
CSS Placement (CSSP)						
$CSSP_{1.5}$			-1.60***	0.07	-1.60***	0.07
$CSSP_2$			-2.83***	0.02	-2.63***	0.02
$CSSP_{2.5}$			-3.82***	0.06	-3.82***	0.06
$CSSP_3$			-5.52***	0.03	-5.52***	0.03
-2 Log L	157,743.00		108,213.72		108,213.74	
AIC	157,751.00		108,231.72		108,229.74	
CCR			71.40%		71.40%	
Score's Test			2,320.02 (df = 15)***		2,240.00 (df = 12)***	

$*p < 5\%$, $**p < 1\%$, $***p < 0.1\%$

Table 7: Three-player cumulative logit model estimates.

5 Application study

In this section, the application study consist of analyzing real bowling data. The data were collected from a local bowling alley in Doha. No names were collected during the data collection. A total of 209 two-player, 211 three-player, and 210 four-player games were observed. Each game contained the rolling scores and the final scores for the CSS. The data were manually entered into a Microsoft Excel spreadsheet. From there, the final scores under the NSS as well as the placings under both systems were calculated.

Number of Players	Scoring Systems CSS	NSS	Both
2	1.44%	5.74%	0.00%
3	5.21%	9.95%	1.42%
4	6.19%	20.48%	1.43%

Table 8: Percentage of observed ties in placings for the application data.

Table 8 shows the percentage of placement ties under the two scoring systems for all two-player, three-player, and four-player games. Of the 209 two-player games, 3 (1.44%) of them resulted in tied placings under the CSS only, 12 (5.74%) of them contained tied games under the NSS only, and none under both systems. For the three-player games, 11 (5.21%) of them contained ties under the CSS only, 21 (9.95%) for the NSS, and 3 (1.42%) under both. Regarding the four-player games, 13 (6.19%) of them contained ties under the CSS, 43 (20.48%) for the NSS, and 3 (1.43%) for both systems.

The applied data involved the collection of 630 bowling games (209 two-player, 211 three-player, and 210 four-player) played by real people at a bowling alley in Doha. The final scores were computed for both scoring systems. Spearman correlations for the two-player, three-player, and four-player games were 0.7811, 0.7495, and 0.7374 respectively. All three correlations were lower than those observed from their respective equal and unequal cases located in Table 4. This could mean that applied data contained more levels of variability than what was simulated.

Placement	Number of Players		
	Two	Three	Four
1	75.73%	65.53%	66.50%
1.5	5.83%	2.43%	2.43%
2	18.45%	25.73%	20.39%
2.5	-	1.94%	2.91%
3	-	4.37%	4.37%
3.5	-	-	1.46%
4	-	-	1.94%

Table 9: Percentages for NSS placements given 1st place for CSS.

Figure 9 contains the percentage of games with specific placements under the NSS conditional on placing first under the CSS. The results in this table can be compared to the distributions found in Table 5. For example, of the two-player games in which there was a bowler that placed first under the CSS, 75.73% of them came in first place under the NSS, 5.83% of them tied for first and second places, and the rest of them (18.45%) placed second under the NSS. Similar to Table 5, bowlers had the highest chance of placing first under the NSS when they placed first under the CSS regardless of the number of players (75.73% for two players, 65.53% for three players, and 66.50% for four players). It is surprising to observe that the percentage for the four-player case was closest to its equal and unequal percentages from Table 5 than the two-player and three-player cases. There is one possible reason for this behavior. Two player and three-player games tend to consist of friends with possible equal ALs. Therefore, each player is just as likely to win as to lose. On the other hand, four-player games tend to consist of families of which there is a father, a mother, and maybe two children. The parent could be viewed as the strongest players based on strength, with the father being the strongest of the two and the children would be the weakest players. Therefore, the AL are most likely unequal. As a result, this could explain why the percentage for the four-player case is closer to the results observed from the simulations. The percentages conditional on placing second are the same as before.

Estimates	Two-Player		Three-Player		Four-Player	
	$\hat{\beta}$	$SE_{\hat{\beta}}$	$\hat{\beta}$	$SE_{\hat{\beta}}$	$\hat{\beta}$	$SE_{\hat{\beta}}$
Intercept ($\hat{\alpha}$)						
$\hat{\alpha}_1$	1.14		0.65		0.65	
$\hat{\alpha}_{1.5}$	1.48		0.75		0.77	
$\hat{\alpha}_2$			2.61		2.33	
$\hat{\alpha}_{2.5}$			3.05		2.54	
$\hat{\alpha}_3$					3.95	
$\hat{\alpha}_{3.5}$					4.39	
CSS Placement (CSSP)						
$CSSP_{1.5}$	-1.31	0.80	-1.19*	0.60	-0.77	0.68
$CSSP_2$	-2.62***	0.23	-1.62***	0.20	-1.95***	0.20
$CSSP_{2.5}$			-2.50***	0.46	-2.85***	0.51
$CSSP_3$			-3.87***	0.24	-3.11***	0.22
$CSSP_{3.5}$					-3.41***	0.59
$CSSP_4$					-4.69***	0.24
-2 Log L	576.27		1,346.70		2,312.49	
AIC	584.27		1,362.70		2,336.49	
CCR	75.36%		60.19%		51.92%	
Score's Test	7.13 (df = 2)*		411.39 (df = 12)***		636.23 (df = 30)***	

$*p < 5\%, **p < 1\%, ***p < 0.1\%$

Table 10: Cumulative logit model estimates for application data.

Table 10 contains the model estimates for the cumulative logit model for the applied data. The deviance and AIC statistics for all three models were higher than the ones from the models created by the simulated data. The statistical significance from CSSP was similar across the three models with the applied data and the models from the simulated data. The CSSP had a significant negative effect on NSS placements. This means that the predicted probability for the 'best' placements (i.e. 1st) decreased as CSSP placement increased, while the predicted probability for the 'worst' placements (i.e. 4th) increased as CSSP placement increased. There are some interesting differences, however, with the results for CSSP. First, the slope effects for CSSP at 1.5 were not statistically significant from the two-player and four-player models. This means that the predicted probabilities for the best NSS placements are not significantly different for players that come in first versus players that tie for first and second under the CSS. Another distinction is that the standard errors for the applied case are seven to eleven times more than the standard errors from the simulated case. There are some possible reasons for this. One reason for the increase could be due to the absence of the AL in the model. Thus, the CSSP standard errors contains the variability behavior from the AL. Another possible reason is the difference in sample sizes between the simulated and applied data have a ratio of 50:1. Therefore, it is trivial to observe standard errors that are much higher for the applied case. The CCR values for the three models ranged between 51.92% for the four-player model to 75.36% with the two-player case. These rates were much lower than the CCR values from the simulated models. Also, the three applied models violated the Score's Test which tests the proportional odds assumption. Despite the test concluding that a possible violation occurred, further analysis showed that the violation did not actually occur.

CSSP	**NSS Placement**						
	1	1.5	2	2.5	3	3.5	4
Two-Player							
1st	75.76%	5.70%	18.54%				
2nd	18.54%	5.70%	75.76%				
Three-Player							
1st	65.77%	2.18%	25.23%	2.30%	4.52%		
2nd	27.60%	2.01%	43.42%	7.69%	19.27%		
3rd	3.87%	0.38%	17.98%	8.42%	69.34%		
Four-Player							
1st	65.86%	2.54%	22.71%	1.59%	5.40%	0.67%	1.23%
2nd	21.48%	2.01%	35.75%	5.06%	23.70%	3.93%	8.08%
3rd	7.91%	0.88%	22.53%	4.79%	33.59%	8.43%	21.87%
4th	1.74%	0.21%	6.65%	1.84%	21.74%	10.25%	57.56%

Table 11: Predicted probabilities for NSS placements for applied data.

Table 11 show the predicted probabilities for placements under the NSS given the CSSP for the applied data. Similar to the Figures 4 tied placements for the CSS were removed although they can be calculated. For example, the predicted probability of placing first under the NSS given a bowler placed first under the CSS is 75.76%. However, the predicted probabilities for a tie for first and second or for placing second are 5.70% and 18.54% respectively.

Similar to the simulation case, the predicted probabilities given that CSSP is second place are complements of the first-place scenario. Note that the probability for maintaining first place under both systems is smaller compared to what was observed from the simulated case. Nevertheless, the predicted probabilities show that bowlers have the highest chances of keeping the same placements under both scoring systems.

6 Conclusions

In this research, researchers investigated how well a NSS, centered upon using powers of ten to keep score, compared to the currently-used scoring system used in recreational and competitive bowling. In comparing the two systems, the purpose was to quantify the consistency of the two systems and determine how often they agree.

The high correlations between the final scores of the two systems suggested that the CSS and NSS are nearly congruent. With regards to the placements, the percentage of times that the two systems agree depend on the number of players involved. The CSS and NSS tended to yield the same placements at least 85% when two players compete, at least 66% of the time when three players compete, and at least 43% of the time when four players compete regardless of the ALs of the players. These percentages increased by at most 4.4%, 11.98%, and 12.20% when one considers partially-concordances respectively. The pattern observed in these percentages suggested that the minimum percentage tend to decrease as more players play together since more players allow for more possible ties in placements and more possible placements to consider. When applied to real data, however, these percentages may be somewhat overestimated.

Modeling the behavior of the placements of the NSS conditional on AL behavior and CSSP is a complicated task, but not impossible. Using the cumulative logit model is one way to show the relationship between these factors, but one should carefully monitor whether the proportional odds assumption is satisfied as it was often violated in the models of this study.

There were several limitations of this study, and these limitations can suggest ideas for future studies. The simulations were based on the use of the binomial distribution but some other discrete distributions could be considered. Considering unequal ALs allowed for different values of p to be used. Also, the values of p increase constantly by 5% between the first and second rolls, and for more skilled bowlers. Future researchers can consider using different increases in this probability or a more complicated distribution to make this behavior more realistic. Also, this study focused on comparing the NSS to the CSS while assuming that the NSS is easier to understand. Other researchers may consider performing a psychological study to determine whether the NSS is easier to understand than the CSS. One may discover that the NSS is better to use for bowlers of particular ages, gender, interest in bowling, or different ALs. Furthermore, a software application could be implemented for the NSS to be commercialized.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

All authors have contributed to all parts of the article. All authors read and approved the final manuscript.

References

Bowling. (n.d.). In *The Columbia Encyclopedia (6th edition)*. Retrieved March 11, 2017 from *Encyclopedia.com* at http://www.encyclopedia.com/reference/encyclopedias-almanacs-transcripts-and-maps/bowling.

Cooper, C. N., & Kennedy, R. E. (1990). A generating function for the distribution of the scores of all possible bowling games. *Mathematics Magazine*, *63*(4), 239-243.

Harmon, J. E. (1985). Bowling Regions of North America. *Journal of Cultural Geography*, *6*(1), 109-124.

Hohn, J. L. (2009). *Generalized Probabilistic Bowling Distributions* (Unpublished master's thesis). Western Kentucky University, Bowling Green, Kentucky.

Keogh, S., & O'Neill, D. (2011). A Statistical Analysis of 10-pin Bowling Scores and an Examination of the Fairness of Alternative Handicapping Systems. *NUI Maynooth* eprint.

Martin, D. E. (2006). Hot-hand effects in sports and a recursive method of computing probabilities for streaks. *Computers & Operations Research*, *33*(7), 1983-2001.

Mullen, M. (2004). *Bowling fundamentals*. Human Kinetics.

Neal, D. K. (2003). Random bowling. *Journal of Recreational Mathematics*, *32*(2), 125-128.

Pluckhahn, Bruce; Bowling games people play. *Bowler's Journal Magazine*, December 1988 issue, 121.

Pretsell, J. M. (1908). *The Game of Bowls: Past and Present*. Oliver and Boyd.

Rogers, M. (2014). Bowling rolls out new scoring system to attract Olympic interest. *USA Today Sports*.

Ryan, C. (2016). Five types of bowling and where to play them. From *Groupon Guide* online magazine. Retrieved from *https://www.groupon.com/articles/bowling-its-many-types-and-where-to-play-them-al*.

Yaari G & David G (2012) "Hot Hand" on Strike: Bowling Data Indicates Correlation to Recent Past Results, Not Causality. PLoS ONE *7*(1): e30112.doi:10.1371/journal.pone.0030112.