Data Driven Decisions, or How to Detect Cheating in Chess

It was once thought chess was a pristine sport, devoid of cheating. After all, the whole chess universe consists of 64 squares, 32 pieces and, as one commentator stated, two players whose heads are eighteen inches apart. Sadly, the use of computers has brought cheating to a whole new level.

To be sure, cheating has existed before computers. Look at the famous claims Bobby Fisher made about international chess in the early 1960s. He claimed the Soviet players colluded to make sure one of them would win the 1962 Candidates tournament. A short video about that follows the text. While the chess world howled at the time that Fisher was nothing but a sore loser, data analysis demonstrated he was correct. The Soviet players played short draws against each other, conserving their energy and strength to defeat players in the tournament from other counties. The chess world came to realize Fisher was correct when the line score (players records against each other) was seriously analyzed. The conclusion was obvious if you looked at the data.

Computers have introduced a whole method of cheating into chess. The chess world has tried to adapt, but skilled cheating has outpaced the detection. Games used to be adjourned after 40 moves because players were exhausted. The game would continue the next day. Players would analyze the position overnight to determine what moves they should make when continuing playing the next day. Soviet players would have an army of seconds who would analyze the position and then give the player the proper set of moves for when the game continued. Even that advantage became moot, as a moderately powered computer could provide the correct strategy in a matter of seconds. Players are now required to complete the game after it begins, no matter how long it takes and how exhausted they become.

A chess players' ability is demonstrated by their Elo rating. It would be too long to go into the explanation of Elo ratings, but a quick example is warranted. The elite players of the world today hover around a rating of 2800 Elo. A player with a 100 point advantage over his opponent would be expected to score 64% of the available points (a win counts as one point, a draw and a half point, and a loss as zero.) Suppose that was a 500 point advantage? The expected number of points to be gathered by the stronger player jumps to 94.7%. In other words, for every 100 games played, the stronger player would be expected to win about 90 games and draw 10, giving them a score of about 95%. Chess AIs such as Alpha Zero have ratings of over 3500 and their power increases every year.

A chess player recently won four games in a row against opponents, each of whom had about a 500 point Elo rating advantage against her. Without going into the gory mathematical computation, and allowing for the fact that draws complicate the computation, the odds of that happening are about 1 in 1.5 million. What was the accusation? Somehow the player was receiving tips from a computer outside of the playing area. The games were being broadcast as they were occurring, so an accomplice could have been relaying computer generated moves to

the alleged cheater. However, there was no direct evidence of this, only the circumstantial evidence generated by the Elo rating.

In both the Fisher case and the recent cheating case, there was no direct evidence to support a cheating accusation Yet, a data driven analysis would lead to a pretty compelling conclusion in both cases. A jury, using that data and using elementary statistics would conclude that cheating had occurred beyond any reasonable doubt.

There are two lessons here:

1. We often have to put aside our beliefs and our intuition and follow the data.
2. Direct evidence may not be available in certain situations. Rigorous analysis of the available data can often lead to virtual certainty and better decisions; and finally,
3. A healthy skepticism is an asset in the business world.