

## How Pingic Rating System Works

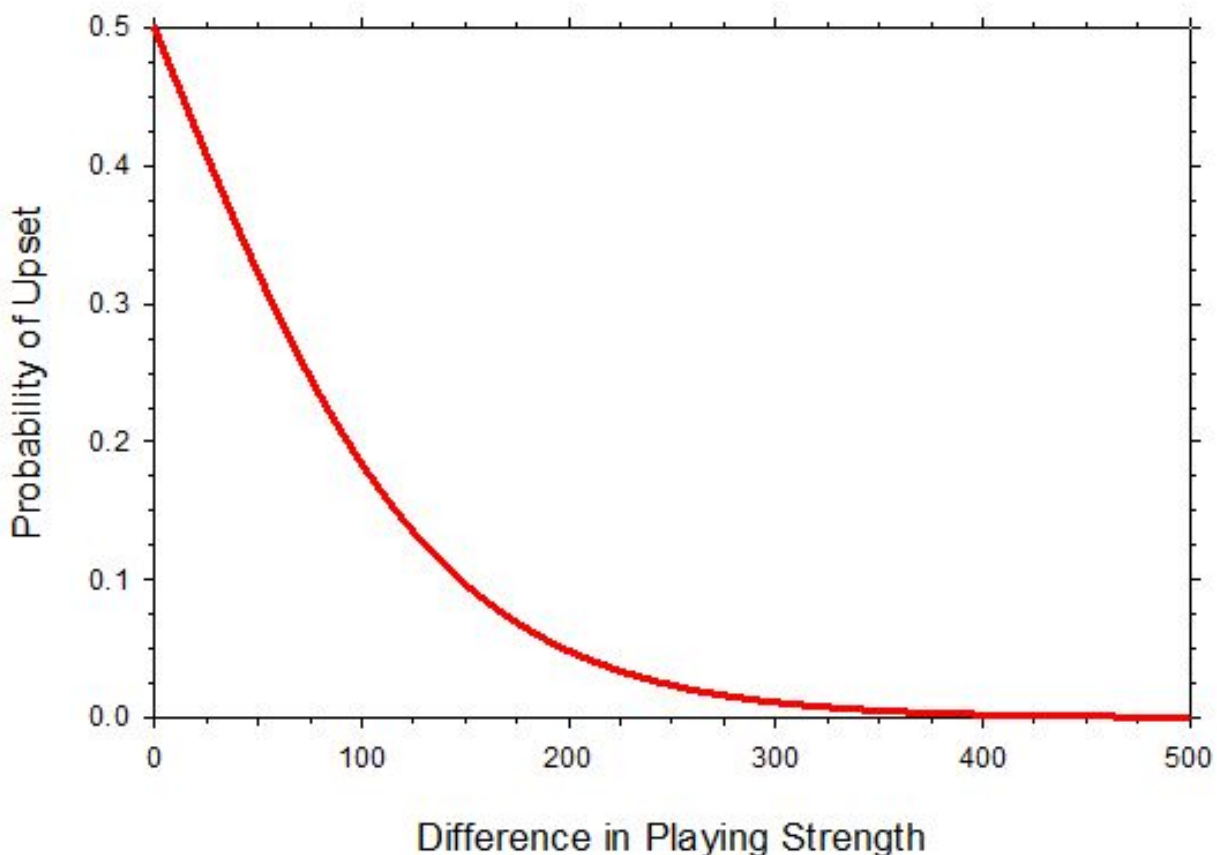
The rating system uses standard concepts and techniques from probability and statistics. This makes the rating system quite different from (and more accurate than) almost all other rating and ranking systems. The following is a nontechnical explanation of how the rating system works.

### Basic Concepts

We presume that each player has a playing strength, i.e., a number that quantifies how strong the player is. The playing strength of a player does not change during a single event, but may change over time, as the player gets better or worse. (An event is a collection of matches, e.g., a tournament, that an event director submits to Ratings Central as a group.)

Even if we knew the playing strengths of two players, we would not know for certain which player would win, since a weaker player will sometimes beat a stronger player. A match is an upset if the player with the lower playing strength wins. We presume that the probability that a match will be an upset is determined solely by the difference in playing strengths of the two players. The larger the difference in the playing strengths, the more likely it is that the stronger player will win. The probability-of-upset function quantifies this.

Probability-of-Upset Function



There are two kinds of probability. There is the probability that one player will defeat another. (This probability is determined by the players' playing strengths.) There is also the probability that a player's playing strength is a certain value (e.g., 1106). The first probability is a property of the players, while the second probability is a property of the rating system.

## Laws

The rating system does not know the playing strengths of players. It only sees match results. The rating system keeps track of what it knows about each player by constructing a law to describe the player's playing strength.

A law is a probability distribution. The rating system assigns a law to each player. The player's law describes the rating system's knowledge of the playing strength of the player. This knowledge is derived from all the match results. The player's law changes with every match that the rating system processes (because the rating system's knowledge of the playing strength of the player changes with every match). From the law, we may determine the probability that the player's playing strength is a certain value (e.g., 1106).

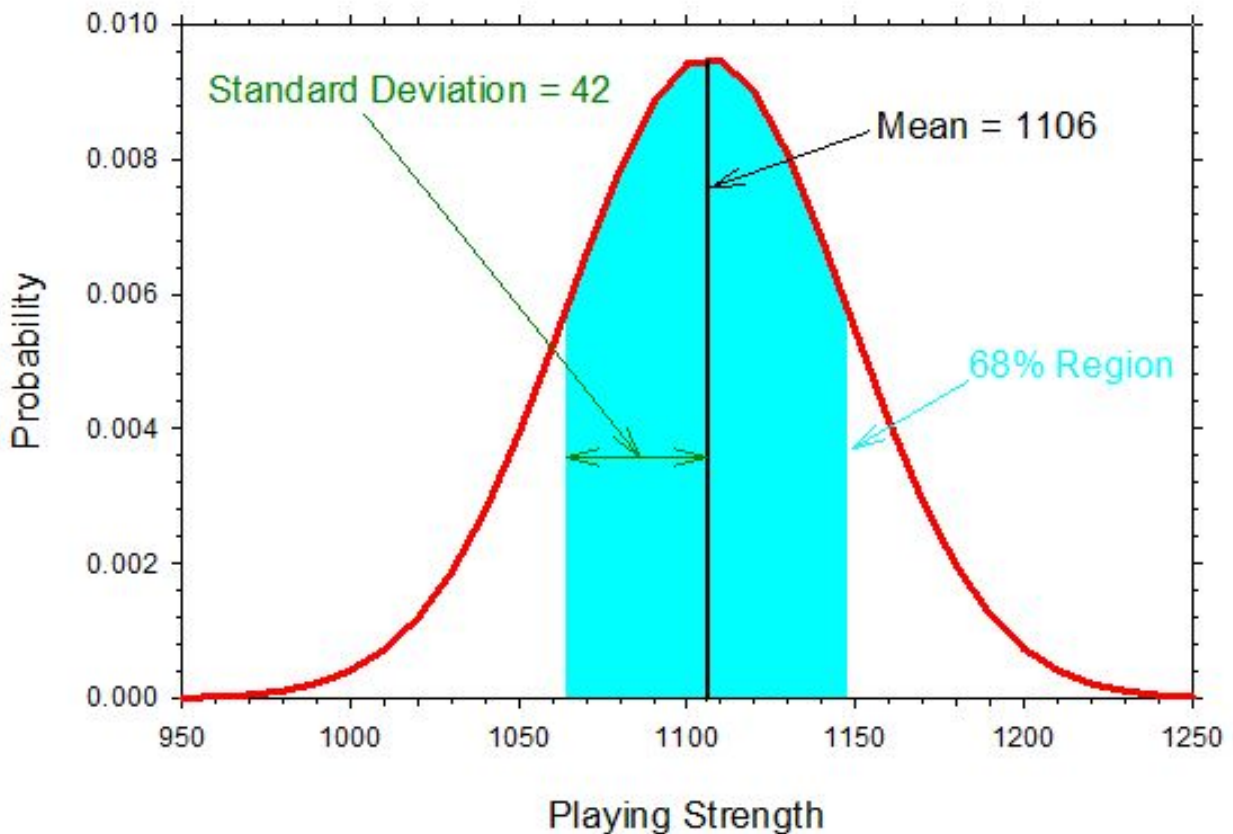
The mean of a law is essentially the location of the center of the law. The mean of a player's law is the rating system's best estimate of the player's playing strength (because it is the center of the rating system's knowledge of the player's playing strength). The mean of a player's law is the rating that the rating system outputs for the player.

The standard deviation measures the spread (width) of a law. The greater the standard deviation of a player's law, the less certain the rating system is of the player's playing strength. The probability that a player's playing strength is within one standard deviation of the mean of the player's law is approximately 68%. The probability that it is within two standard deviations is approximately 95%. The probability that it is within three standard deviations is approximately 99.7%.

If the meaning of a sentence like "The probability that a player's playing strength is within two standard deviations is 95%" isn't clear, here is another way of saying the same thing: There is a 95% probability that the player's playing strength is between the mean minus twice the standard deviation and the mean plus twice the standard deviation. For example, if the mean is 1106 and the standard deviation is 42, then

- There is a 68% probability that the player's playing strength is between 1064 and 1148.
- There is a 95% probability that the player's playing strength is between 1022 and 1190.
- There is a 99.7% probability that the player's playing strength is between 980 and 1232.

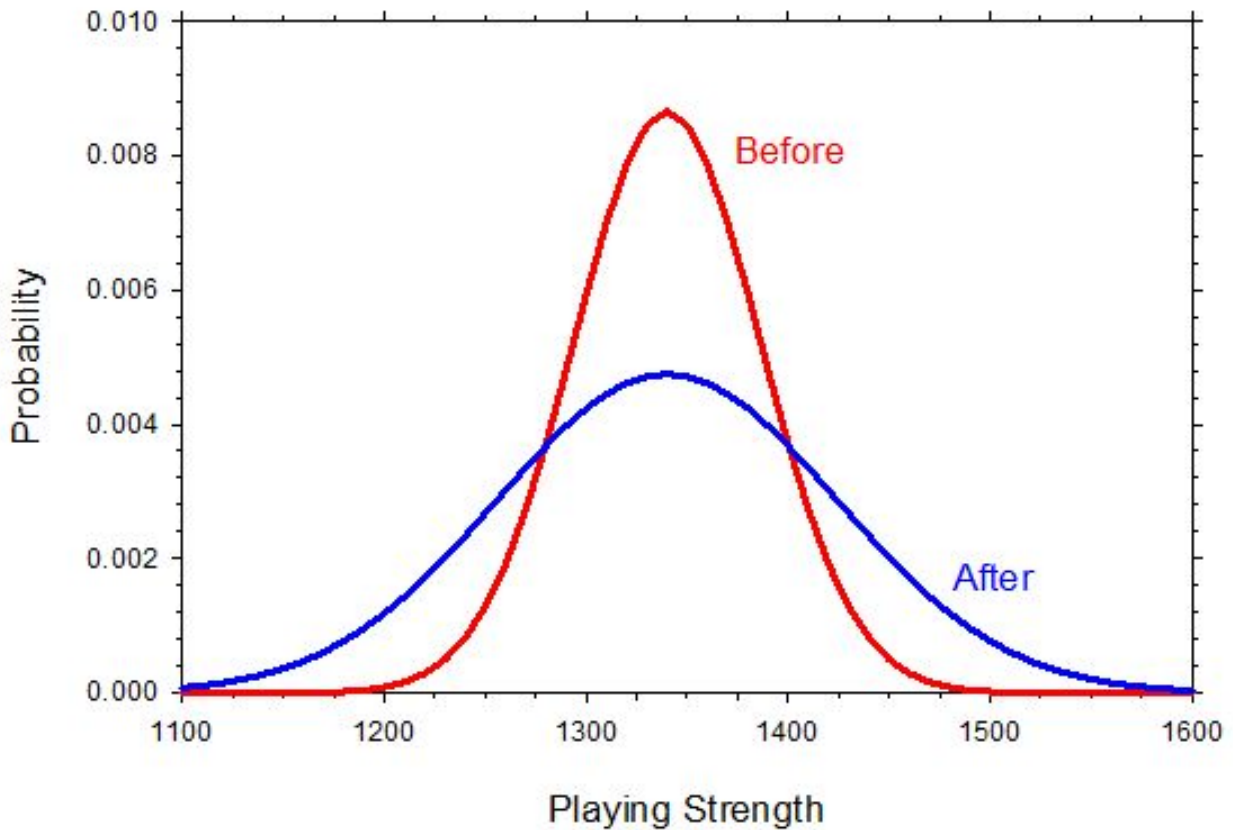
## Typical Law



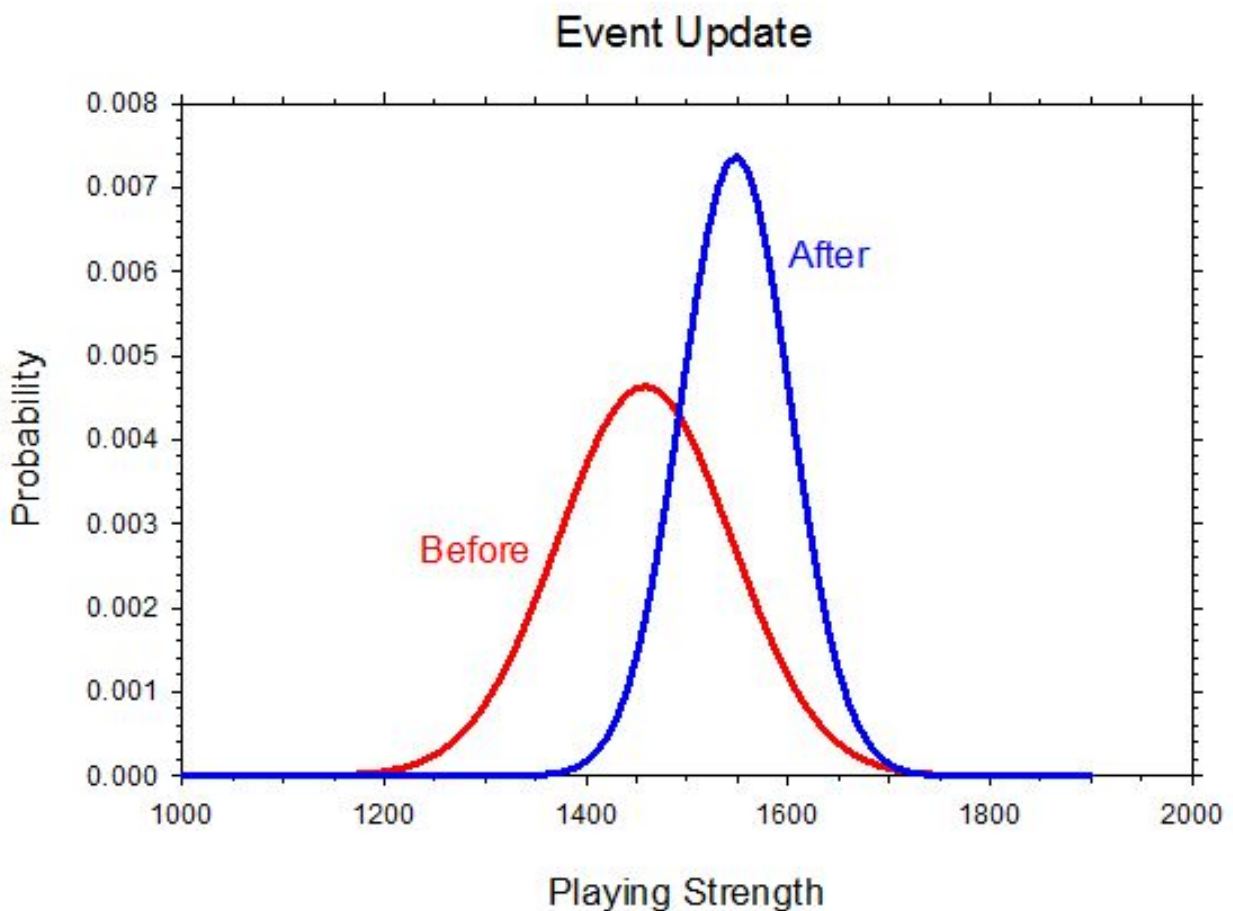
## Updates

A player's playing strength may change over time, as the player gets better or worse. Because of this, the more time that has passed since the last event that the player played, the less certain we are of the player's playing strength. The process of updating a player's law to take into account the passage of time is the temporal update. The temporal update makes a law more spread out, but doesn't change the mean. For example, if the rating system has not seen a player for a year, the standard deviation of the player's law will increase by at most 79.4 points and the mean will increase by 7 points. However, the increase in mean will not be credited until the player plays in an event. (The new temporal-update is the sum of a mean-zero normal random walk with a variance of  $70^2$  per year and a Poisson jump process with jumps of 200 points, a mean of 7 points per year, and a variance of  $200 \times 7$  per year.)

## Temporal Update



The process of updating a player's law to take into account the player's results in a single event is the event update. In theory, we should process all the results of an event as a single group. However, we need something that is computationally feasible. When doing the event update, the rating system only looks at a player's results and the results of each of the player's opponents. This is similar to what you might do at a tournament: Suppose that you lose a match and think your opponent is better than their rating. The way you might check this is to look at the draw sheets to see how your opponent did against other players. These are the same results that the rating system looks at when updating your rating.



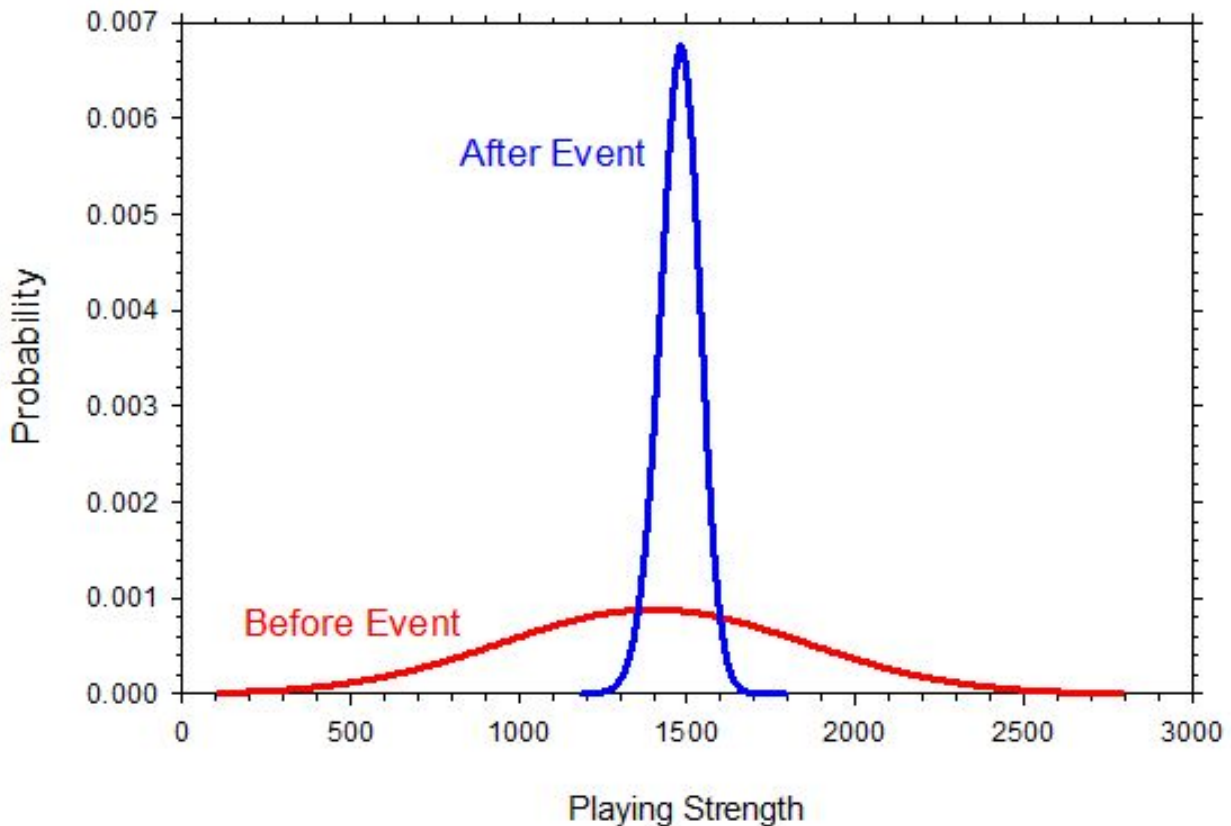
A small standard deviation makes it harder both to gain and to lose points. If a player with a small standard deviation plays a player with a large standard deviation, then the former's rating will change less than the latter's.

The rating system only cares about who wins a match, not what the score is.

## New Players

The rating system assigns a law to each new (unrated) player. This law reflects what we expect the playing strength of a new player in the particular event to be. Event directors tell us what mean and standard deviation we should use for their event based on their experience. Event directors can also specify a law for a particular new player based on their knowledge of that player. Often, the standard deviation for a new player will be large reflecting the range in playing strength of unrated players that may enter the event. After playing an event or two, the standard deviation of a new player's law should drop significantly. How quickly this happens depends on how many matches the player plays, the outcome of the matches, and the laws of the player's opponents.

## Typical Laws for New Player



## Event Processing

Here are the steps that the rating system goes through when processing an event:

1. Assign a law to each unrated player.
2. Retrieve the law of each rated player from the database and apply the temporal update.
3. For each player:
  1. Calculate an adjusted law (defined below) for each of the player's opponents.
  2. Update the player's law for each of the player's matches using in each case the adjusted law for the player's opponent.

The adjusted law is the opponent's law updated for all of the opponent's matches except for the matches with the current player. The adjusted law depends on both the player and the opponent. So, the same opponent will have different adjusted laws when different players are being processed.

## References

- Marcus, D. J. (2001) [New Table-Tennis Rating System](#). Journal of the Royal Statistical Society: Series D (The Statistician), 50: 191-208. doi: 10.1111/1467-9884.00271
- Marcus, D. J. (2011a) [Ratings Central: Accurate, Automated, Bayesian Table Tennis Ratings for Clubs, Leagues, Tournaments, and Organizations](#). Joint Statistical Meetings, July 30-August 4, 2011.
- Marcus, D. J. (2011b) [Ratings Central: Accurate, Automated, Bayesian Table Tennis Ratings for Clubs, Leagues, Tournaments, and Organizations](#). NESSIS (New England Symposium on Statistics in Sports), September 24, 2011.