

2009 年臺灣國際科學展覽會

優勝作品專輯

國家：United States of America

編號：010041

作品名稱

Baseball and the Markov Chain Theory

得獎獎項

Mathematics Second Award

作者姓名

Kevin Evgene Fritz

Baseball and the Markov Chain Theory

Abstract

In contrast to an analysis based on simple probabilities of events, a Markov Chain analysis takes into account the influence that the current state of a process has on future states. In this project, I have applied a Markov Chain analysis to model probabilities associated with the game of baseball.

For this research an existing algorithm developed by Bukiet et al. was modified and improved. Play-by-play data for an entire season of Major League baseball were acquired and a computer program was written to parse this data to determine the probabilities of baserunners advancing a specific number of bases on various types of hits. These probabilities were used to create an improved probability model for runner advancement which was included in the algorithm. The effect of double plays on runner advancement probabilities was also determined and implemented in the algorithm. A series of computer programs in the C language were written to automate implementation of the model across many players or teams. Finally, the main computer program was modified to add the capability of analyzing partial games to find the effect of certain plays on the probability of winning. The new Markov Chain model was applied to five research questions.

The study identified the incremental gain in bases advanced for various types of hits relative to a single. The probability of the pitcher's team winning when a great slugger was intentionally walked was found to be lower than when the slugger was pitched to in most game situations.

Regardless of the inning or batting position used, National League teams score more runs when using a pinch hitter and both the inning and batting position had an effect on the number of runs scored. Players were rated according to the number of wins they would add to the average team on the basis of this Markov model and the results were compared to actual winners of MVP and Cy Young awards. Sportswriters selected the highest rated player for the MVP and Cy Young awards only 36% and 50% of the time, respectively.

評語

Among the high school mathematics curricula in Taiwan, Statistics is one topic that most teachers find it difficult to teach and students find it hard to learn. What causes this problem? Here is one explanation: the topic is not treated concretely. Like in the United States, Baseball here is so popular that a large population is glued to the TV set whenever there is a game being played. Therefore, exploring Baseball Statistics is a perfectly concrete learning activity that develops the true feeling for Statistics.