

Game on: Sports Analytics Conference produces a growing avalanche of quantitative research

March 2 2012, by Peter Dizikes



Thousands attended last year's Sloan Sports Analytics Conference. Photo: Justin Jensen

Imagine if your hobby became your profession — or, indeed, if it started a whole new profession.

This year, the MIT Sloan [Sports](#) Analytics Conference features someone who can make that claim: Bill James, the baseball analyst whose pioneering work opened up the entire field. On Friday and Saturday, James will speak for the first time at the MIT conference, the biggest event of its kind.

A high school English teacher and factory watchman from Kansas, in 1977 James began self-publishing an annual baseball book analyzing how well the sport's conventional wisdom held up to hard statistical evidence. By 1982, “The Bill James Baseball Abstract” had gained a national

publisher and wide following.

For his readers at the time, James' work was a revelation, providing the statistical tools that let fans grasp baseball better than some of the sport's executives did. By 2002, James even became a baseball insider himself, as an adviser to the Boston Red Sox.

Many of the 2,000 attendees at this year's conference, held at Boston's Hynes Convention Center, will also be insiders: Professional teams now routinely employ analytics staffers, and Houston Rockets general manager Daryl Morey MBA '00 is an MIT Sloan graduate who co-founded the conference.

Yet sports analytics is still open to anyone who wants to give it a try. This year's conference has a burgeoning [research-paper competition](#) in which outsiders shed light on sports. *MIT News* spoke to a few MIT-affiliated entrants.

Big data tips the pitch

Baseball teams employ elaborate signs to prevent hitters from knowing what kind of pitch is coming next. But what if hitters could figure it out anyway? A new paper by Gartheeban Ganeshapillai, a PhD candidate in MIT's Computer Science and Artificial Intelligence Laboratory (CSAIL), and John Guttag, the Dugald C. Jackson Professor of Computer Science and Engineering, uses machine-learning techniques to show that some pitchers are highly predictable — once you know what to look for.

In their paper, "Predicting the Next Pitch," Gartheeban and Guttag constructed a database of every single pitch delivered in major league baseball in 2008 and 2009, then correlated each hurler's pitch selection with dozens of variables, such as the score and the count on the hitter.

They found that the most telling predictors of pitch selection are the pitcher's history against the batter, and the pitcher's tendency to throw certain pitches in certain counts — such as a fastball with a 2-2 count.

“The count on its own was not that influential, but if you look at a particular count for a particular pitcher, that is very relevant,” Gartheeban says. Compared to a model that emphasizes how frequently a pitcher throws fastballs, Gartheeban and Guttag found that their method produces a mean fastball-prediction improvement of more than 18 percent, and an improvement of as much as 311 percent for some pitchers.

However, Guttag notes, “We can't say yet if falling into a pattern is a good thing or a bad thing.” One highly predictable pitcher in the study is the New York Yankees' famous reliever Mariano Rivera — who relies heavily on one pitch, his cut fastball, which is simply very hard to hit squarely.

As it happens, Guttag is a lifelong Yankees fan who [sometimes wears the team's cap on the MIT campus](#). “I grew up watching Mickey Mantle and Roger Maris and Whitey Ford at the old Yankee Stadium,” says Guttag, a native of the New York area. Gartheeban grew up in Sri Lanka and has attended only one major league ballgame, but loves cricket, which he regards as overdue for serious sports-analytics work.

More typically, the duo analyzes medical data. But Guttag thinks this project was a sufficiently challenging technical exercise, and says he may assign similar problems to incoming graduate students.

Chemistry lessons

Basketball announcers love to talk about “chemistry,” by which they usually mean the mix of personalities on a team. But Robert Ayer MBA

'11, a management consultant in Atlanta, has found a more analytical way of seeing if players will blend well.

In his conference paper, Ayer examined which combinations of individual skills best complement each other in a team context. “How players fit together is a key part of winning, but it hasn’t been quantified in a serious way,” he says.

Ayer’s research involves a two-step analysis. First, by using data since 1977 and looking at 14 different individual measures — such as shots attempted and turnovers per game — Ayer constructed a set of 13 types of NBA players, based on the way their statistical profiles clustered together. These include point guards who love to shoot (Type 7), or power forwards who stretch defenses by taking outside shots (Type 11).

Secondly, Ayer developed a regression model to extract which combinations of players, other things being equal, produce the best results. He found that his Type 8 player —multiskilled wing players who can score from outside, but also have a lot of assists, like the Boston Celtics’ Paul Pierce — tend to wind up on successful teams.

“There are not a whole lot of multifaceted small forwards out there, but they will fit well with almost anyone you put around them,” Ayer says. “Teams with one of those guys will almost always overperform their individual talent and coaching.” On the other hand, Ayer found, teams in which both starting guards shoot a lot usually underperform.

This analysis could help general managers select players. When using the NBA draft to complement a Type 8 player, “You can pick just the best player available,” says Ayer, who despite his time at MIT remains an Atlanta Hawks fan. “You wouldn’t have to reach for a guy who you think is not as talented but fits a certain role.” On the other hand, he notes, “Conventional wisdom is that duplicative players don’t fit well, and for

the most part, that was confirmed by the research.”

Splash! When teams go in the tank

Veteran basketball observers have long suspected that some bad teams “tank,” or stop trying, in order to improve their chances in the NBA draft lottery. Now Christopher Walters and Tyler Williams, PhD candidates in MIT’s Department of Economics, have the numbers to prove it.

Every year, 14 NBA teams qualify for the lottery; the worse a team’s record, the better its chances of getting the top pick, which sometimes yields a superstar along the lines of LeBron James or Dwight Howard.

Examining data since 1985, when the lottery was instituted, Walters and Williams have found that, other things being equal, teams eliminated from the playoffs mysteriously lose 14 percent more frequently when they have the chance to improve their lottery odds than in games where their lottery odds are unaffected.

“I was a little surprised at the magnitude of the effect,” Williams says, adding, “That 14 percent figure is probably the lowest, safest estimate.”

However, teams that seem to tank do so in subtle ways. “It’s not so much about being lackadaisical on the court,” says Williams, a longtime University of Michigan sports fan who normally studies education issues. “Teams will rest their best players if they’re a little injured, or try new rotations of players. And fans generally understand they’re trying to add an exciting player.”

The paper gets a related result — about the value of the top pick — by controlling for what economists call “propensity scores,” something the researchers encountered in their graduate work. “We thought it was a cool application of an idea we’d learned about in our economics

training,” Williams adds. Today’s part-time sports analysts may still be pursuing it as a hobby — but it’s a serious one.

This story is republished courtesy of MIT News (web.mit.edu/newsoffice/), a popular site that covers news about MIT research, innovation and teaching.

Provided by Massachusetts Institute of Technology

Citation: Game on: Sports Analytics Conference produces a growing avalanche of quantitative research (2012, March 2) retrieved 10 April 2024 from <https://phys.org/news/2012-03-game-sports-analytics-conference-avalanche.html>

<p>This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.</p>
