

From the Diamond to the Rink: Applications of Statistics in Sports

Phil Yates

Department of Mathematics
Saint Michael's College
pyates@smcvt.edu

July 31, 2011



Outline

- Brief summary of previous talks

Outline

- Brief summary of previous talks
- How has research in statistics in sports grown

Outline

- Brief summary of previous talks
- How has research in statistics in sports grown
- The future of research of statistics in sports

Outline

- Brief summary of previous talks
- How has research in statistics in sports grown
- The future of research of statistics in sports
- Summary: Statistics and sports – "An all-encompassing discipline"

NHL Player Ability Ratings (A. C. Thomas)

- Improve upon \pm ratings with actual ability ratings

NHL Player Ability Ratings (A. C. Thomas)

- Improve upon \pm ratings with actual ability ratings
- Scoring rates for each team modeled as a semi-Markov process

NHL Player Ability Ratings (A. C. Thomas)

- Improve upon \pm ratings with actual ability ratings
- Scoring rates for each team modeled as a semi-Markov process
- Hazard functions depend on players actually on the ice

NHL Player Ability Ratings (A. C. Thomas)

- Improve upon \pm ratings with actual ability ratings
- Scoring rates for each team modeled as a semi-Markov process
- Hazard functions depend on players actually on the ice
- Stable estimates of player ability are computed

NHL Player Ability Ratings (A. C. Thomas)

- Improve upon \pm ratings with actual ability ratings
- Scoring rates for each team modeled as a semi-Markov process
- Hazard functions depend on players actually on the ice
- Stable estimates of player ability are computed
- Possible interaction effects can be identified to see which pairs of players are better suited playing with one another

A Curveball Index (J. Wilson)

- No system to rate breaking balls

A Curveball Index (J. Wilson)

- No system to rate breaking balls
- Assign measurements to curve ball based on 5 aspect of a good curveball

A Curveball Index (J. Wilson)

- No system to rate breaking balls
- Assign measurements to curve ball based on 5 aspect of a good curveball
- Data obtained from 30 pitches (3 pitchers, 10 pitches each) rated on a 0 to 100 scale by pitching coach

A Curveball Index (J. Wilson)

- No system to rate breaking balls
- Assign measurements to curve ball based on 5 aspect of a good curveball
- Data obtained from 30 pitches (3 pitchers, 10 pitches each) rated on a 0 to 100 scale by pitching coach
- Data were different than what can be obtained from Pitch F/X

A Curveball Index (J. Wilson)

- No system to rate breaking balls
- Assign measurements to curve ball based on 5 aspect of a good curveball
- Data obtained from 30 pitches (3 pitchers, 10 pitches each) rated on a 0 to 100 scale by pitching coach
- Data were different than what can be obtained from Pitch F/X
- Regression was valid with meaningful coefficients and no multicollinearity

A Curveball Index (J. Wilson)

- No system to rate breaking balls
- Assign measurements to curve ball based on 5 aspect of a good curveball
- Data obtained from 30 pitches (3 pitchers, 10 pitches each) rated on a 0 to 100 scale by pitching coach
- Data were different than what can be obtained from Pitch F/X
- Regression was valid with meaningful coefficients and no multicollinearity
- Future work: More data and maybe analysis of additional pitch (slider)

Analysis of the NHL Draft (M. Schuckers)

- Analysis on players selected in the NHL Entry Drafts from 1988 to 1997 (www.hockey-reference.com, www.nhl.com, & www.eliteprospects.com)

Analysis of the NHL Draft (M. Schuckers)

- Analysis on players selected in the NHL Entry Drafts from 1988 to 1997 (www.hockey-reference.com, www.nhl.com, & www.eliteprospects.com)
- Method similar to NFL Draft Pick Value Chart developed by Schuckers (*JQAS*, 2010)

Analysis of the NHL Draft (M. Schuckers)

- Analysis on players selected in the NHL Entry Drafts from 1988 to 1997 (www.hockey-reference.com, www.nhl.com, & www.eliteprospects.com)
- Method similar to NFL Draft Pick Value Chart developed by Schuckers (*JQAS*, 2010)
- Look at number of games played by position over NHL career (season is 82 games long)

Analysis of the NHL Draft (M. Schuckers)

- Analysis on players selected in the NHL Entry Drafts from 1988 to 1997 (www.hockey-reference.com, www.nhl.com, & www.eliteprospects.com)
- Method similar to NFL Draft Pick Value Chart developed by Schuckers (*JQAS*, 2010)
- Look at number of games played by position over NHL career (season is 82 games long)
- NHL Pick Value Draft Chart splined together two non-parametric regression curves to regress games played onto selection

Analysis of the NHL Draft (M. Schuckers)

- Analysis on players selected in the NHL Entry Drafts from 1988 to 1997 (www.hockey-reference.com, www.nhl.com, & www.eliteprospects.com)
- Method similar to NFL Draft Pick Value Chart developed by Schuckers (*JQAS*, 2010)
- Look at number of games played by position over NHL career (season is 82 games long)
- NHL Pick Value Draft Chart splined together two non-parametric regression curves to regress games played onto selection
- Chart assigns value to each player based on the expected number of career games that the player will spend in the NHL

Analysis of NHL On-Ice Officials (L. C. Brozowski)

- Four officials on the ice – two referees and two linesmen

Analysis of NHL On-Ice Officials (L. C. Brozowski)

- Four officials on the ice – two referees and two linesmen
- Goal: identify which teams are more likely to commit penalties, which officials are more likely to be on ice when penalty is called, and does the score impact whether penalty is called

Analysis of NHL On-Ice Officials (L. C. Brozowski)

- Four officials on the ice – two referees and two linesmen
- Goal: identify which teams are more likely to commit penalties, which officials are more likely to be on ice when penalty is called, and does the score impact whether penalty is called
- Logistic regression was used with penalty as a “success” based on play-by-play data from www.nhl.com

Analysis of NHL On-Ice Officials (L. C. Brozowski)

- Four officials on the ice – two referees and two linesmen
- Goal: identify which teams are more likely to commit penalties, which officials are more likely to be on ice when penalty is called, and does the score impact whether penalty is called
- Logistic regression was used with penalty as a “success” based on play-by-play data from www.nhl.com
- Two officials were on ice for a significant larger amount of penalties

Analysis of NHL On-Ice Officials (L. C. Brozowski)

- Four officials on the ice – two referees and two linesmen
- Goal: identify which teams are more likely to commit penalties, which officials are more likely to be on ice when penalty is called, and does the score impact whether penalty is called
- Logistic regression was used with penalty as a “success” based on play-by-play data from www.nhl.com
- Two officials were on ice for a significant larger amount of penalties
- More penalties occur when teams are up by two or three goals

Analysis of NHL On-Ice Officials (L. C. Brozowski)

- Four officials on the ice – two referees and two linesmen
- Goal: identify which teams are more likely to commit penalties, which officials are more likely to be on ice when penalty is called, and does the score impact whether penalty is called
- Logistic regression was used with penalty as a “success” based on play-by-play data from www.nhl.com
- Two officials were on ice for a significant larger amount of penalties
- More penalties occur when teams are up by two or three goals
- Future work: Look at more than 2009-2010 season and spatial analysis to see where penalties occur the most on the ice

Baseball & Hockey

Baseball

- *The American Statistician*: 119 articles through 2005

Baseball & Hockey

Baseball

- *The American Statistician*: 119 articles through 2005
 - First mention of baseball in article - 1949
 - First article about baseball - 1969 (incidence of double headers)

Baseball & Hockey

Baseball

- *The American Statistician*: 119 articles through 2005
 - First mention of baseball in article - 1949
 - First article about baseball - 1969 (incidence of double headers)
- *JASA*: 100 articles through 2005

Baseball & Hockey

Baseball

- *The American Statistician*: 119 articles through 2005
 - First mention of baseball in article - 1949
 - First article about baseball - 1969 (incidence of double headers)
- *JASA*: 100 articles through 2005
 - First mention of baseball in article - 1923
 - First article about baseball - 1961 (progress of score in game)

Baseball & Hockey

Baseball

- *The American Statistician*: 119 articles through 2005
 - First mention of baseball in article - 1949
 - First article about baseball - 1969 (incidence of double headers)
- *JASA*: 100 articles through 2005
 - First mention of baseball in article - 1923
 - First article about baseball - 1961 (progress of score in game)
- *JQAS*: 100 articles (2005 to 2011)

Baseball & Hockey

Baseball

- *The American Statistician*: 119 articles through 2005
 - First mention of baseball in article - 1949
 - First article about baseball - 1969 (incidence of double headers)
- *JASA*: 100 articles through 2005
 - First mention of baseball in article - 1923
 - First article about baseball - 1961 (progress of score in game)
- *JQAS*: 100 articles (2005 to 2011)
- SABR - Society for American Baseball Research (founded 1971, first publication 1972)

Baseball & Hockey

Hockey

- *The American Statistician*: 28 articles through 2005

Baseball & Hockey

Hockey

- *The American Statistician*: 28 articles through 2005
 - First article about hockey - 1977 (Poisson & goal scoring)

Baseball & Hockey

Hockey

- *The American Statistician*: 28 articles through 2005
 - First article about hockey - 1977 (Poisson & goal scoring)
- *JASA*: 23 articles through 2005

Baseball & Hockey

Hockey

- *The American Statistician*: 28 articles through 2005
 - First article about hockey - 1977 (Poisson & goal scoring)
- *JASA*: 23 articles through 2005
 - First article about hockey - 1966 (Best k of $2k - 1$ comparisons)

Baseball & Hockey

Hockey

- *The American Statistician*: 28 articles through 2005
 - First article about hockey - 1977 (Poisson & goal scoring)
- *JASA*: 23 articles through 2005
 - First article about hockey - 1966 (Best k of $2k - 1$ comparisons)
- *JQAS*: 46 articles (2005 to 2011)

Data Sources

Baseball ● www.baseball-reference.com

Data Sources

Baseball

- `www.baseball-reference.com`
- `www.retrosheet.org`

Data Sources

Baseball

- www.baseball-reference.com
- www.retrosheet.org
- www.fangraphs.com

Data Sources

Baseball

- www.baseball-reference.com
- www.retrosheet.org
- www.fangraphs.com
- Major League Baseball Advanced Baseball - Pitchf/x data

Data Sources

Baseball

- www.baseball-reference.com
- www.retrosheet.org
- www.fangraphs.com
- Major League Baseball Advanced Baseball - Pitchf/x data

Hockey

- www.hockey-reference.com

Data Sources

Baseball

- www.baseball-reference.com
- www.retrosheet.org
- www.fangraphs.com
- Major League Baseball Advanced Baseball - Pitchf/x data

Hockey

- www.hockey-reference.com
- www.nhl.com

Data Sources

Baseball

- www.baseball-reference.com
- www.retrosheet.org
- www.fangraphs.com
- Major League Baseball Advanced Baseball - Pitchf/x data

Hockey

- www.hockey-reference.com
- www.nhl.com
- www.hockeydb.com

Baseball & Hockey

Baseball

- “There is nothing on earth anybody can do with fielding.” - Branch Rickey, 1954

Baseball & Hockey

Baseball

- "There is nothing on earth anybody can do with fielding." - Branch Rickey, 1954
- Pitching metrics – FIP (Fielding Independent Pitching), xFIP (Expected Fielding Independent Pitching), BABIP (Batting Average on Balls in Play)

Baseball & Hockey

Baseball

- "There is nothing on earth anybody can do with fielding." - Branch Rickey, 1954
- Pitching metrics – FIP (Fielding Independent Pitching), xFIP (Expected Fielding Independent Pitching), BABIP (Batting Average on Balls in Play)
- **An All-Encompassing Discipline:** Joint work with trainers, doctors, and statisticians to evaluate injuries

Hockey

- Fangraphs for hockey?

Baseball & Hockey

Baseball

- "There is nothing on earth anybody can do with fielding." - Branch Rickey, 1954
- Pitching metrics – FIP (Fielding Independent Pitching), xFIP (Expected Fielding Independent Pitching), BABIP (Batting Average on Balls in Play)
- **An All-Encompassing Discipline:** Joint work with trainers, doctors, and statisticians to evaluate injuries

Hockey

- Fangraphs for hockey?
- PAR (points above replacement) - how much \$ on free agent market?

Baseball & Hockey

Baseball

- "There is nothing on earth anybody can do with fielding." - Branch Rickey, 1954
- Pitching metrics – FIP (Fielding Independent Pitching), xFIP (Expected Fielding Independent Pitching), BABIP (Batting Average on Balls in Play)
- **An All-Encompassing Discipline:** Joint work with trainers, doctors, and statisticians to evaluate injuries

Hockey

- Fangraphs for hockey?
- PAR (points above replacement) - how much \$ on free agent market?
- Line combinations

Thank You!

Organizer Vittorio Addona

Chair Patricia English

Speakers

- Andrew Thomas
- Jason Wilson
- Michael Schuckers
- Lauren Brozowski