Analysis and Prediction of the 2001 Indians Baseball Season

Andrew Chernichenko
Electrical and Computer Engineering
Portland State University

Outline
- Introduction
- Objectives
- Why is it important
- Gathering and analysis of the data
- What was found
- Conclusion

Introduction
- Cleveland Indians
- Data was gathered primarily from websites and mainly from espn.com and indians.com
- Analysis of the winning percentage vs. loosing percentage
- Data analyzed: 1999, 2000 and 2001 seasons

Objectives
- Prediction and outcome of the 2001 season and individual games.
- My financial standing at the end of this baseball season.

Significance
- Important for owners and managers of the team.
- Betting purposes.
- Personal gain.

Analysis
- Analysis of the previous year- 1999.
- Graph showing winning percentage over games played.
Analysis

• Analysis of the previous year- 2000.

Analysis

• Analysis of the 2001 season.

Results

<table>
<thead>
<tr>
<th>Year</th>
<th>Win. Pct.</th>
<th>Opp. Pct.</th>
<th>Mean</th>
<th>σ</th>
</tr>
</thead>
<tbody>
<tr>
<td>1999</td>
<td>0.599</td>
<td>0.473</td>
<td>0.657</td>
<td>0.083</td>
</tr>
<tr>
<td>2000</td>
<td>0.556</td>
<td>0.501</td>
<td>0.576</td>
<td>0.076</td>
</tr>
<tr>
<td>2001</td>
<td>0.644</td>
<td>0.465</td>
<td>0.615</td>
<td>0.072</td>
</tr>
</tbody>
</table>

Prediction

• Used MATLAB with 1999 season as input
• Two weights on the data
  – Opponent
  – Injuries
Prediction

Conclusion

• 1999 and 2001 seasons are similar
• More games played-less variance
• Individual players do influence the outcome of the game.
• Opponents capabilities correlate to wins/losses