Understanding probability using games
Outline for today

Better know a player

Review of concepts in probability

Multiplication rule

Strat-o-matic
Announcement: class final projects

Begin putting together your class final project

Should focus on research question
  • Doesn’t have to be about baseball, but need to find data set that you can use to answer the question

1-2 paragraph proposal is due the Wednesday March 29th

Projects presentation are on May 3rd – good to start working on them soon!
Better know a player
Probability

Probability is a way of measuring the uncertainty of the outcome of an event

Q: The set of all possible random outcomes is called the…?
   • A: Sample space

A probability distribution Pr(X) assigns a value between 0 and 1 to all events
   • (an event is a subset of the sample space)
If a fair ball is hit (1 or 6 rolled first) then the following table indicates the outcome of the play:

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single</td>
<td>Out</td>
<td>Out</td>
<td>Out</td>
<td>Out</td>
<td>Error</td>
</tr>
<tr>
<td>Out</td>
<td>Double</td>
<td>Single</td>
<td>Out</td>
<td>Single</td>
<td>Out</td>
</tr>
<tr>
<td>Out</td>
<td>Single</td>
<td>Triple</td>
<td>Out</td>
<td>Out</td>
<td>Out</td>
</tr>
<tr>
<td>Out</td>
<td>Out</td>
<td>Out</td>
<td>Out</td>
<td>Out</td>
<td>Out</td>
</tr>
<tr>
<td>Out</td>
<td>Single</td>
<td>Out</td>
<td>Out</td>
<td>Out</td>
<td>Single</td>
</tr>
<tr>
<td>Error</td>
<td>Out</td>
<td>Out</td>
<td>Out</td>
<td>Single</td>
<td>Home run</td>
</tr>
</tbody>
</table>

Q: What is the sample space there?
A: \{single, double, triple, home run, out, error\}
Big League baseball probability distribution if a ball is hit in play

Q: why is this a valid probability distribution?
A: All values are between 0 and 1 and the sum is 1
Probability rules - Additive rule

If there are two events A, and B, then the probability of A or B happening is:

\[ \Pr(A \text{ or } B) = \Pr(A) + \Pr(B) - \Pr(A, B) \]

Events are called **mutually exclusive** if events A and B can not both occur

- i.e., \( \Pr(A, B) = 0 \)
Q: If the ball is in play, what is the probability of getting a hit in big league baseball?

Pr(hit) =

= Pr(1B or 2B or 3B or HR)

= Pr(1B) + Pr(2B) + Pr(3B) + Pr(HR)

= 7/36 + 1/36 + 1/36 + 1/36

= 10/36
Multiplicative Rule

\[ \Pr(A, B) = \Pr(A | B) \times \Pr(B) \]

E.g., Suppose we draw 2 cards from a 52 card deck. What is the probability they are both diamonds?

\[ \Pr(D_1, D_2) = \Pr(D_1) \times \Pr(D_2 | D_1) \]
\[ = \frac{13}{52} \times \frac{12}{51} \]
\[ = 0.12 \]
Special Case: Multiplicative Rule for independent events

Two events are independent if:

$$\Pr(A, B) = \Pr(A) \times \Pr(B)$$

Q: what is the probability of getting two strikes in a row in Big League baseball?

$$\Pr(\text{Strike, Strike}) = \Pr(\text{Strike}) \times \Pr(\text{Strike})$$

$$\begin{align*}
&= \frac{1}{3} \times \frac{1}{3} \\
&= \frac{1}{9}
\end{align*}$$
For any one pitch (not assuming that the play is in play), what is the probability of the following events?

- A home run?
- A out?
- A single?
- A Hit?

<table>
<thead>
<tr>
<th>1st Die</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Single</td>
<td>Out</td>
<td>Out</td>
<td>Out</td>
<td>Out</td>
<td>Error</td>
</tr>
<tr>
<td>2</td>
<td>Out</td>
<td>Double</td>
<td>Single</td>
<td>Out</td>
<td>Single</td>
<td>Out</td>
</tr>
<tr>
<td>3</td>
<td>Out</td>
<td>Single</td>
<td>Triple</td>
<td>Out</td>
<td>Out</td>
<td>Out</td>
</tr>
<tr>
<td>4</td>
<td>Out</td>
<td>Out</td>
<td>Out</td>
<td>Out</td>
<td>Out</td>
<td>Out</td>
</tr>
<tr>
<td>5</td>
<td>Out</td>
<td>Single</td>
<td>Out</td>
<td>Out</td>
<td>Out</td>
<td>Single</td>
</tr>
<tr>
<td>6</td>
<td>Error</td>
<td>Out</td>
<td>Out</td>
<td>Out</td>
<td>Single</td>
<td>Home run</td>
</tr>
</tbody>
</table>
**Big League Baseball**

**For any one pitch** (not assuming that the play is in play), what is the probability of the following events?

- A home run? $\frac{1}{36} \times \frac{1}{3}$
- A out? $\frac{24}{36} \times \frac{1}{3}$
- A single? $\frac{7}{36} \times \frac{1}{3}$
- A Hit? $\frac{10}{36} \times \frac{1}{3}$

<table>
<thead>
<tr>
<th>1st Die</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Single</td>
<td>Out</td>
<td>Out</td>
<td>Out</td>
<td>Out</td>
<td>Error</td>
</tr>
<tr>
<td>2</td>
<td>Out</td>
<td>Double</td>
<td>Single</td>
<td>Out</td>
<td>Single</td>
<td>Out</td>
</tr>
<tr>
<td>3</td>
<td>Out</td>
<td>Single</td>
<td>Triple</td>
<td>Out</td>
<td>Out</td>
<td>Out</td>
</tr>
<tr>
<td>4</td>
<td>Out</td>
<td>Out</td>
<td>Out</td>
<td>Out</td>
<td>Out</td>
<td>Out</td>
</tr>
<tr>
<td>5</td>
<td>Out</td>
<td>Single</td>
<td>Out</td>
<td>Out</td>
<td>Out</td>
<td>Single</td>
</tr>
<tr>
<td>6</td>
<td>Error</td>
<td>Out</td>
<td>Out</td>
<td>Out</td>
<td>Single</td>
<td>Home run</td>
</tr>
</tbody>
</table>
Comparisons with observed data

How do the theoretical proportions compared to the empirical proportions for balls in play for:

- Hit rate, error rate, single rate, and home run rate?

<table>
<thead>
<tr>
<th></th>
<th>Observed</th>
<th>Theoretical</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hit rate</td>
<td>0.36</td>
<td>0.28</td>
</tr>
<tr>
<td>Error rate</td>
<td>0.06</td>
<td>0.06</td>
</tr>
<tr>
<td>Single rate</td>
<td>0.18</td>
<td>0.19</td>
</tr>
<tr>
<td>Home run rate</td>
<td>0.05</td>
<td>0.03</td>
</tr>
</tbody>
</table>
All-star baseball

For all-star baseball is it not possible to analytically calculate the probability of different events.

• Instead we would just have to repeat the event many times can calculate the relative frequencies.
Strat-o-matic

Much more complex board games
  • Takes into account Hitters and Pitchers
  • Advanced version accounts for additional factors (e.g., ball parks etc.)
Each player is represented by a card

A single white die is rolled to determine whether to use hitter or pitcher’s card:

- 1-3 -> hitters card
- 4-6 -> pitcher’s card
Then, two red dice are rolled and their sum determines which play in the card should be used.

For some plays additionally a 20 sided die is rolled to determine the final outcome

- and other tables/rules often need to be consulted
Let’s play:

1. Divide into two teams: Dodgers and Cardinals
   • ~9 people on each team

2. Each team member needs to select a position

3. For each position, everyone needs to select a player at that position:
   • Positions listed on top right:
     • Lower numbers means better fielding
     • Also look at hitting statistics
We are going to keep it simple so we will ignore a few rules...

- No base stealing, ignore ++,

- For batter’s card, we can ignore which position the ball was hit to
  - E.g., pay attention to flyball A, ignore (rf)

- For pitcher’s card, when there is an X we will use the position (card) the ball was hit to and refer to a table

Batting will roll all the dice...
  - Except for fielding plays

Play ball!
Worksheet 6

Due on midnight on Sunday March 26th

> source('/home/shared/baseball_stats_2017/baseball_class_functions.R')

> get.worksheet(6)