

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

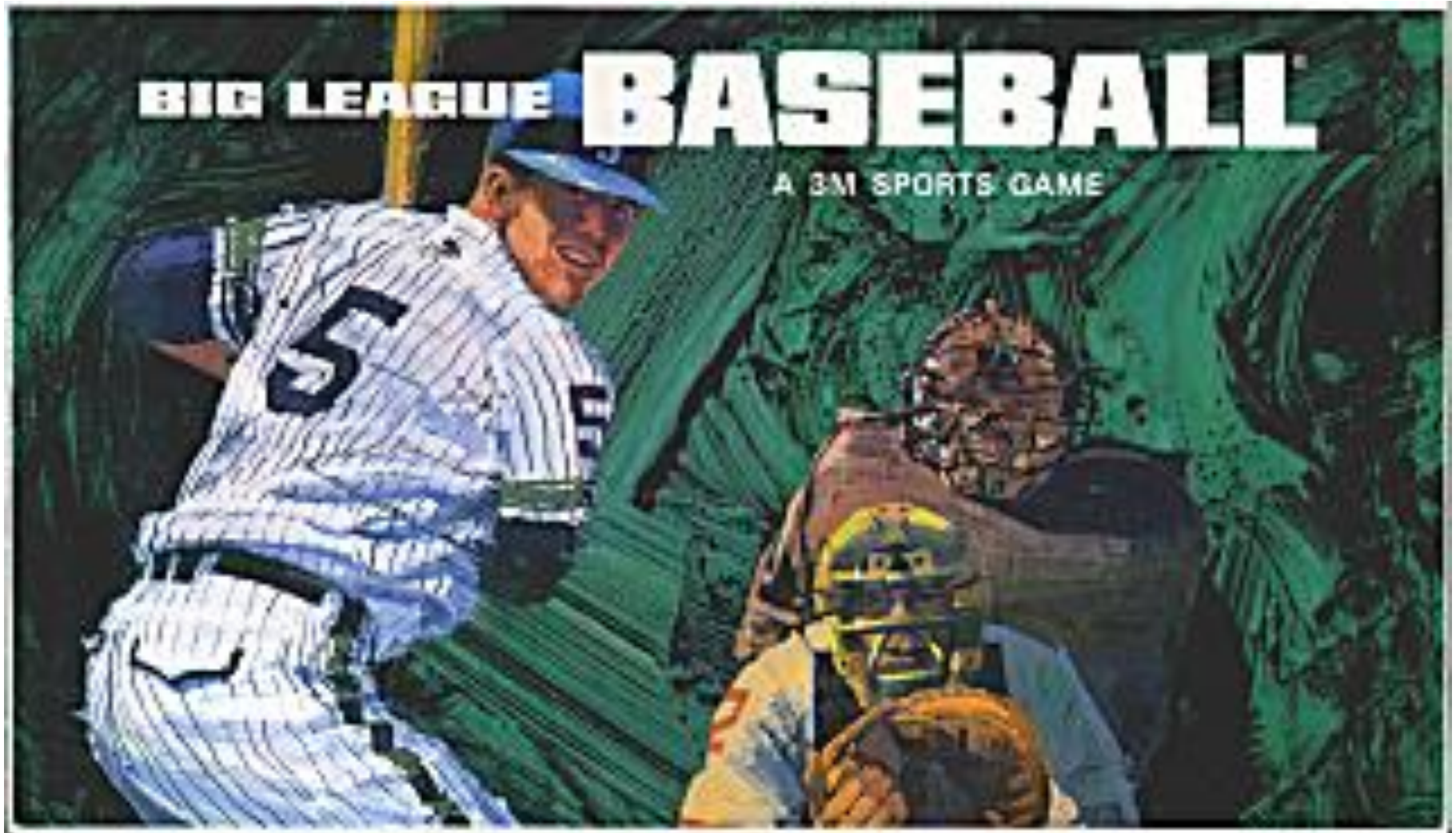
Q: The set of all possible random outcomes is called the...?

- A: **Sample space**

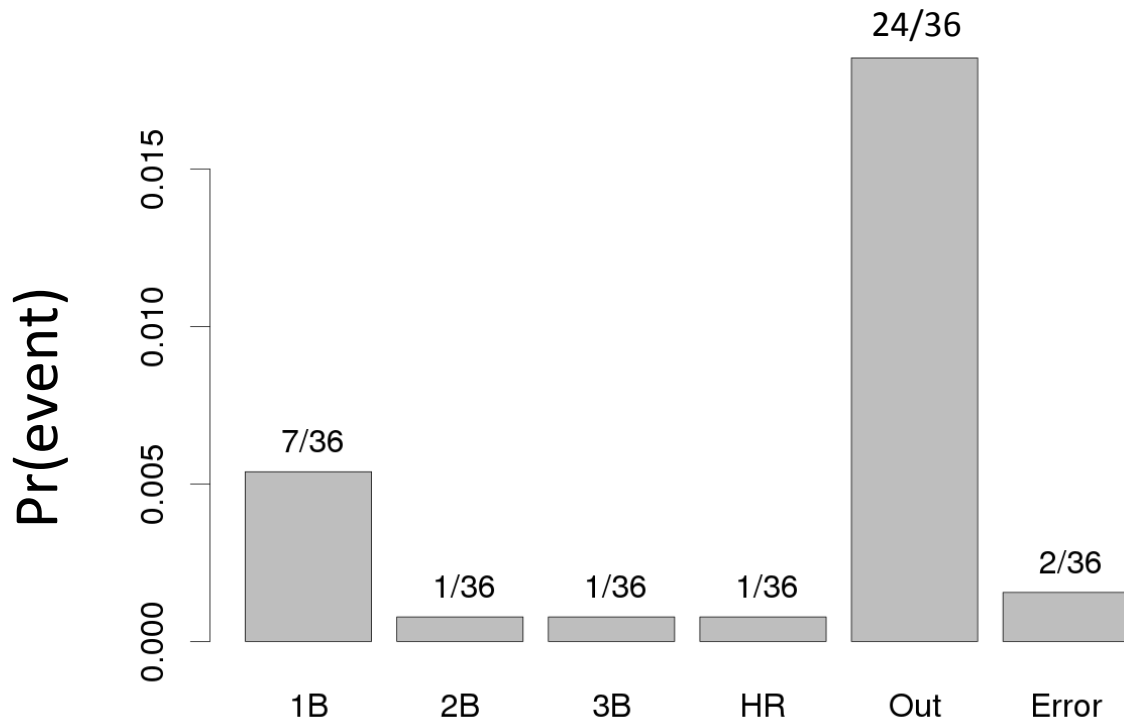
A probability distribution $\Pr(X)$ assigns a value between 0 and 1 to all events

- (an event is a subset of the sample space)

Big League Baseball Probability Distribution



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:

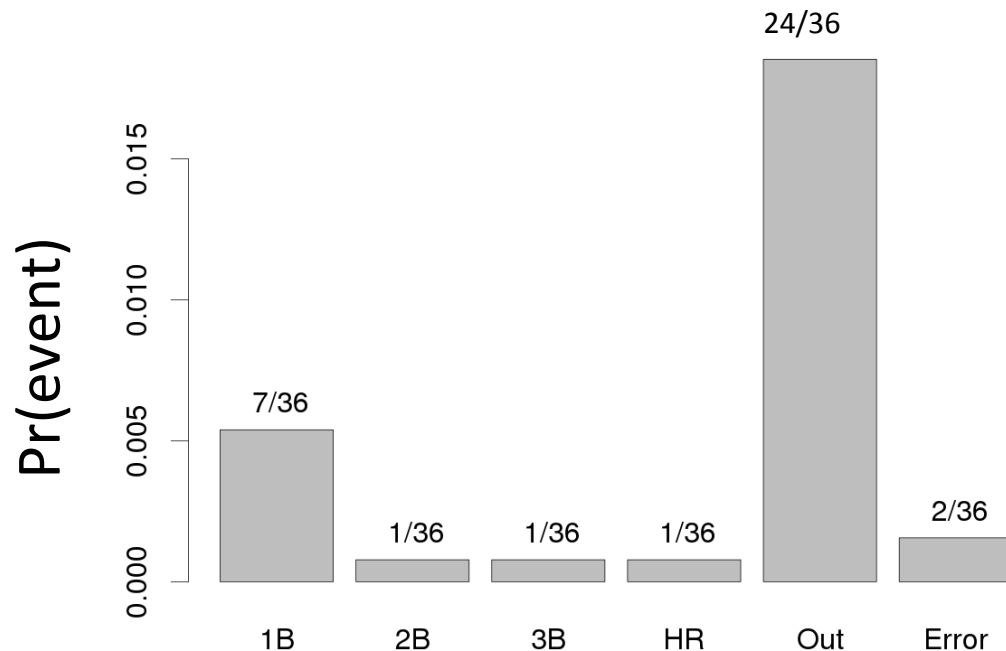
$$\Pr(A \text{ or } B) = \Pr(A) + \Pr(B) - \Pr(A, B)$$

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

- i.e., $\Pr(A, B) = 0$

Probability rules - Additive rule

Q: If the ball is in play, what is the probability of getting a hit in big league baseball?



$\Pr(\text{hit}) =$

$= \Pr(1\text{B or } 2\text{B or } 3\text{B or HR})$

$= \Pr(1\text{B}) + \Pr(2\text{B}) + \Pr(3\text{B}) + \Pr(\text{HR})$

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

$= 10/36$

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?

$$\begin{aligned}\Pr(D1, D2) &= \Pr(D1) \times \Pr(D2 | D1) \\ &= 13/52 \times 12/51 = 0.12\end{aligned}$$

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?

$$\begin{aligned} \Pr(\text{Strike}, \text{Strike}) &= \Pr(\text{Strike}) \times \Pr(\text{Strike}) \\ &= \frac{1}{3} \times \frac{1}{3} = \frac{1}{9} \end{aligned}$$

Big League Baseball

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

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

2nd Die

1st Die

	1	2	3	4	5	6
1	Single	Out	Out	Out	Out	Error
2	Out	Double	Single	Out	Single	Out
3	Out	Single	Triple	Out	Out	Out
4	Out	Out	Out	Out	Out	Out
5	Out	Single	Out	Out	Out	Single
6	Error	Out	Out	Out	Single	Home run

Big League Baseball

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

- A home run? $1/36 * 1/3$
- A out? $24/36 * 1/3$
- A single? $7/36 * 1/3$
- A Hit? $10/36 * 1/3$

2nd Die

1st Die

	1	2	3	4	5	6
1	Single	Out	Out	Out	Out	Error
2	Out	Double	Single	Out	Single	Out
3	Out	Single	Triple	Out	Out	Out
4	Out	Out	Out	Out	Out	Out
5	Out	Single	Out	Out	Out	Single
6	Error	Out	Out	Out	Single	Home run

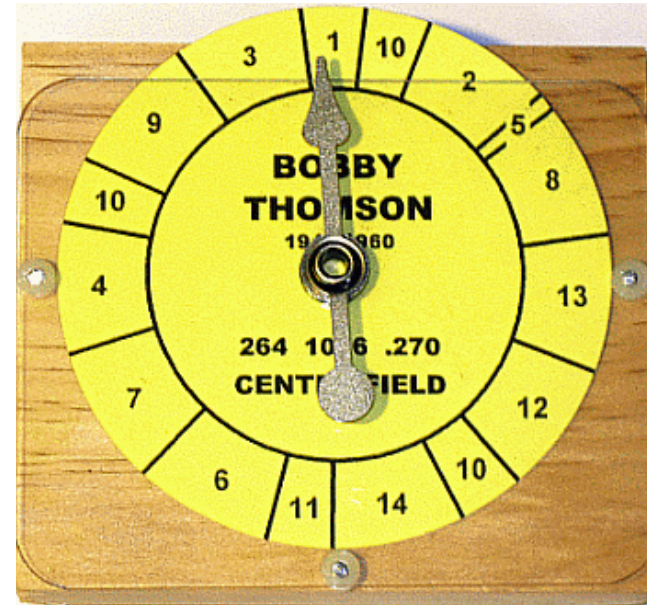
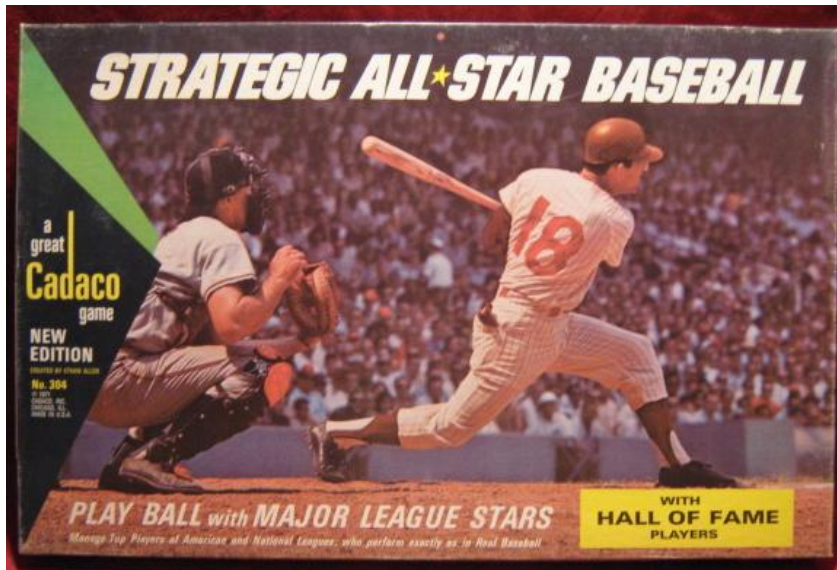
Comparisons with observed data

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?

	Observed	Theoretical
Hit rate	0.36	0.28
Error rate	0.06	0.06
Single rate	0.18	0.19
Home run rate	0.05	0.03

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.)

Strat-o-matic

NELSON CRUZ rightfield-2 stealing-A
leftfield-3 running 1-12

TEXAS

1	2	3
2-flyball (cf)B 3-groundball (3b)A 4-groundball (3b)B 5-strikeout 6-strikeout 7-WALK 8-groundball (ss)A 9-strikeout 10-groundball (p)B 11-groundball (ss)A++ 12-lineout (2b) into as many outs as possible	2-flyball (cf)A 3-groundball (ss)A++ 4-groundball (ss)A 5-strikeout 6-strikeout 7-flyball (lf)B 8-strikeout 9-strikeout 10-flyball (cf)B 11-groundball (ss)A++ plus injury 12-WALK	2-SINGLE* 1-2 lineout (3b) 3-20 3-WALK 4-groundball (3b)B 5-SINGLE 6-HOMERUN 7-HOMERUN 1-10 DOUBLE 11-20 8-groundball (3b)A 9-DOUBLE** 1 SINGLE** 2-20 10-HOMERUN 11-SINGLE 12-flyball (rf)B

2009 BATTING RECORD

AVG	AB	2B	3B	HR	RBI
.260	462	21	1	33	76
BB	SO	SB	CS	SLG%	ONBASE%
49	118	20	.4	.524	.332

BRANDON McCARTHY pitcher-3 starter
batting # 1

PITCHING CARD

TEXAS

4	5	6
2-groundball (p)B 3-FLYBALL (lf)X 4-flyball (cf)B 5-groundball (2b)C 6-SINGLE* 1-10 lineout (3b) 11-20 7-flyball (rf)B 8-DOUBLE** 1-12 SINGLE** 13-20 9-flyball (lf)C 10-popout (1b) 11-FLYBALL (rf)X 12-flyball (lf)B	2-lineout (ss) 3-TRIPLE 1 DOUBLE 2-20 4-CATCHER'S CARD X 5-HOMERUN 1-19 DOUBLE 20 6-strikeout 7-GROUND-BALL(2b)X 8-flyball (cf)B 9-strikeout 10-GROUND-BALL(ss)X 11-strikeout 12-groundball (1b)C	2-groundball (1b)C 3-GROUND-BALL(p)X 4-GROUND-BALL(3b)X 5-strikeout 6-WALK 7-SINGLE 8-WALK 9-GROUND-BALL(ss)X 10-FLYBALL (cf)X 11-GROUND-BALL(1b)X 12-groundball (p)B

2009 PITCHING RECORD

W	L	ERA	STARTS	SAVES
7	4	4.62	17	0
IP	HITS ALLOWED	BB	SO	HOMERUNS ALLOWED
97	96	36	65	13

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

Strat-o-matic

NELSON CRUZ rightfield-2 stealing-A
leftfield-3 running 1-12

TEXAS

1	2	3
2-flyball (cf)B 3-groundball (3b)A 4-groundball (3b)B 5-strikeout 6-strikeout 7-WALK 8-groundball (ss)A 9-strikeout 10-groundball (p)B 11-groundball (ss)A++ 12-lineout (2b) into as many outs as possible	2-flyball (cf)A 3-groundball (ss)A++ 4-groundball (ss)A 5-strikeout 6-strikeout 7-flyball (lf)B 8-strikeout 9-strikeout 10-flyball (cf)B 11-groundball (ss)A++ plus injury 12-WALK	2-SINGLE* 1-2 lineout (3b) 3-20 3-WALK 4-groundball (3b)B 5-SINGLE 6-HOMERUN 7-HOMERUN 1-10 DOUBLE 11-20 8-groundball (3b)A 9-DOUBLE** 1 SINGLE** 2-20 10-HOMERUN 11-SINGLE 12-flyball (rf)B

2009 BATTING RECORD

AVG	AB	2B	3B	HR	RBI
.260	462	21	1	33	76
BB	SO	SB	CS	SLG%	ONBASE%
49	118	20	.4	.524	.332

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

Strat-o-matic

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

Strat-o-matic

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)
```