Math/Stat 341: Probability: Fall '21 (Williams)

Professor Steven J Miller: sjm1@williams.edu

Homepage:
https://web.williams.edu/Mathematics/sjmiller/public_html/341Fa21

Lecture 07: 9-24-21:

Plan for the day: Lecture 07: September 24, 2021:

https://web.williams.edu/Mathematics/sjmiller/public_html/341Fa21/handouts/341Notes_Chap1.pdf

- Work in progress, joint with Professor Cleary of Babson, opportunity for student participation in papers….
- **Key question: What are the right metrics, what is the right data?**

General items.

- How do we compare? How do we model?

Classifying GOATs (like Brady, Russell and Ruth) by Measuring Their Tails.

Rick Cleary, Babson College
Steve Miller, Williams College

rcleary@babson.edu, sjm1@williams.edu

MathFest: August 5, 2021

BIG PICTURE SPORTS QUESTIONS…

-Who is the GOAT (Greatest Of All Time) in a particular sport?

-Who is GOAT of GOATs across sports?

We recognize that this is not a well posed question … but fans and media try to answer it. To do so they make mathematical and statistical arguments that lead them to particular metrics … sometimes without even realizing it!

Examples abound…

- A student in the school paper at Nova Southeastern University makes the case for Tom Brady in football: TomBradyGOATArgument

- Justin Quinn in USAToday says it’s Bill Russell in basketball: BillRussellGOATArgument

- And at Quora.com Mike Berard makes the case for Babe Ruth: BabeRuthGOATArgument

The Brady and Russell arguments are largely team oriented, while

Ruth’s case is more about his individual excellence.

OUR GOAL TODAY…

This is a *preliminary* report on our work to date. We want to:

-Show that metrics matter.

-Give examples of ways that assumptions lead to metrics.

-Choose explicit metrics first and use them to evaluate 'something like' a GOAT argument; maybe a 'best teammate".

Note: The ideas here can be applied in teaching many applications besides sports. Consistent with ideas in social choice, finance, economics and other fields. Can adjust the technical level to be anything from a first year seminar to a capstone project!

METRICS MATTER - Who’s IN first?

How can we tell who the GOAT if we can’t even decide GORN (Greatest of Right Now)!

Consider June 4, 2018. Boston Red Sox were 41-19, winning percentage .683. New York Yankees were 37-17, winning percentage .685.

ESPN *correctly* reports Red Sox in first (HURRAH!) since they are one “game ahead.”

GOOGLE reports Yankees in first (BOOOO!) since they have higher winning percentage.

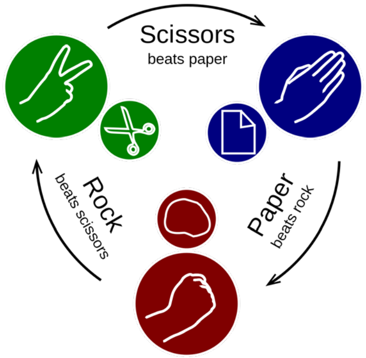
American League

EAST	W	L	PCT	GB
Boston Red Sox	41	19	.683	-
New York Yankees	37	17	.685	1
Tampa Bay Rays	28	30	.483	12
Toronto Blue Jays	26	33	.441	14.5
Baltimore Orioles	17	41	.293	23

American League

AL East	W	L	Pct	GB
Yankees	37	17	.685	1.0
Red Sox	41	19	.683	-
Rays	28	30	.483	12.0
Blue Jays	26	33	.441	14.5
Orioles	17	41	.293	23.0

Or … Cross country example



Order of finish:

A - B - B - A - A - C - C - B - C - C - C - C - C - B - A - B - A - A - B - B

“Invitational” Scoring: A wins!!

A: 1+4+5+15+17 = 42; C: 6+7+9+10+11 = 43, B: 2+3+8+14+16 = 43.   (C gets second since their sixth runner beat B’s sixth runner.)

“Dual Meet” Scoring: Three different races:

B beats A 27-28; A beats C 29-30; and C beats B 28-31! Everybody goes 1-1!

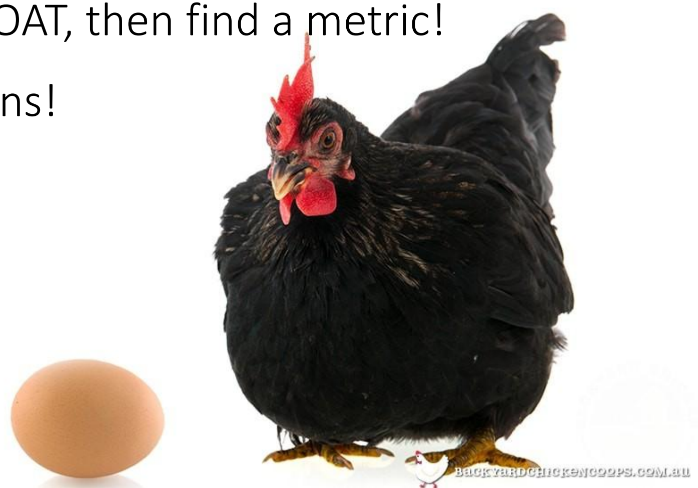
WHICH COMES FIRST? THE METRIC OR THE GOAT?

You don’t simply choose the GOAT … you choose a metric!

Scientific method … Choose the metric and see who’s the GOAT;

OR
As a fan/writer with deadline for a column… Choose the GOAT, then find a metric!

We will pick some reasonable metrics and see what happens!



Let’s have a vote!

In the chat, tell us who you think is the best candidate for GOAT of GOATs:

BRADY … NFL

RUSSELL … NBA

RUTH … MLB

OTHER … Provide Name/Sport

Let's follow up with some data

In the chat, we had votes for the best candidate for GOAT of GOATs:

BRADY … NFL

RUSSELL … NBA

RUTH … MLB

OTHER … Provide Name/Sport

AN ALTERNATE PLACE WE CHOSE TO START: Who is the BOAT (Best Of All Teammates)? Measured by team success relative to league size/quality/playoff format. Preliminary research suggests Brady and Russell is the place to start looking.

The BOAT must … Get to the playoffs!

Under simplest assumptions, with goal of ‘make the playoffs’. Each team equally likely to qualify each year, years are independent, so a binomial model.

Brady: 18 times to playoff in 20 year career in 32 team league with 12 playoff qualifiers. Appearances ~ binomial (20, 12/32).

P(18 or more in 20 years) = $3.02 * 10^{-9}$ = .00000000302.

Brady vs. Russell playoff appearances

Brady: 18 times to playoff in 20 year career in 32 team league with 12 playoff qualifiers. Appearances ~ binomial (20, 12/32).

P(18 or more in 20 years) = $3.02 * 10^{-9}$ = .00000000302.

Russell: Thirteen for thirteen in making playoffs … but in small league where more than half of teams made playoffs!

P(13 for 13 in making playoffs) = $6.29 * 10^{-3}$ = .00629.

A point in Brady's favor here!

But if we do titles instead of appearances:

Under simplest assumptions: Each team equally likely to win each year, years are independent, so a binomial model:

Brady: Seven titles in 20 years in 32 team league.

X ~ binom(20, 1/32)

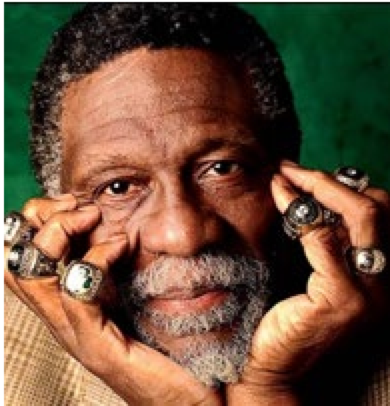
P (X at least 7) = $1.49 * 10^{-6}$ = .00000149.

Russell: Eleven titles in 13 years in (approx.) 10 team league.

X ~ binom(13, 1/10)

P(X at least 11) = $6.44 * 10^{-10}$ = .000000000644.

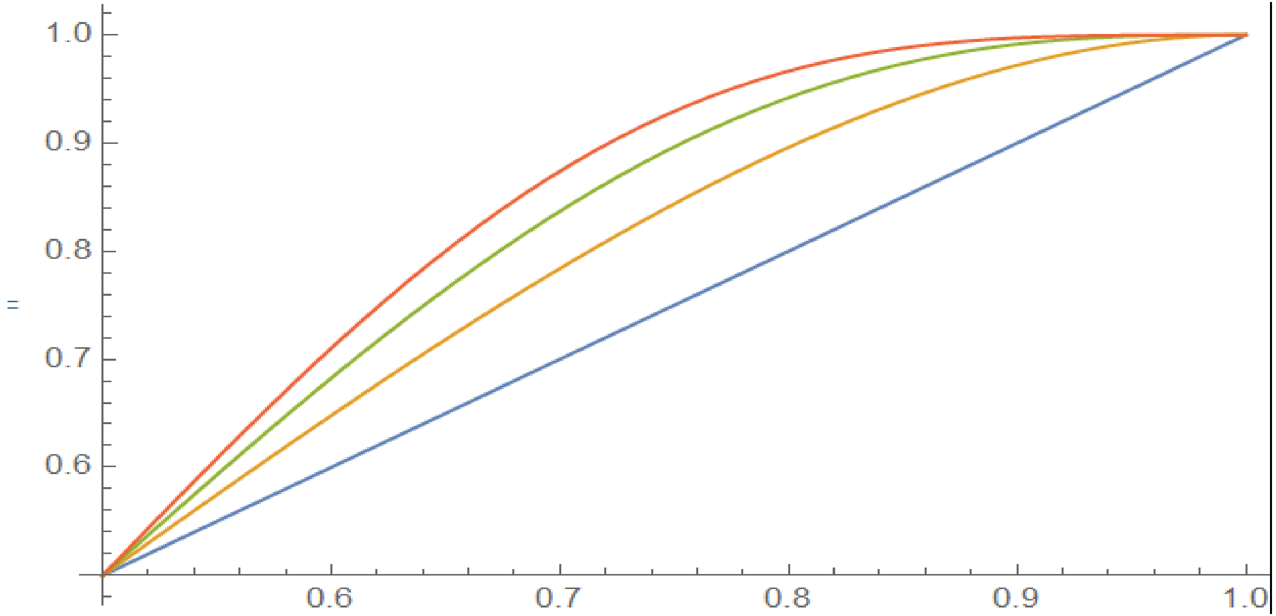
Russell looks better here … again, metrics matter!!



BUT Brady played *GAMES;* Russell *SERIES*

Horizontal axis: P(stronger team wins any one game).

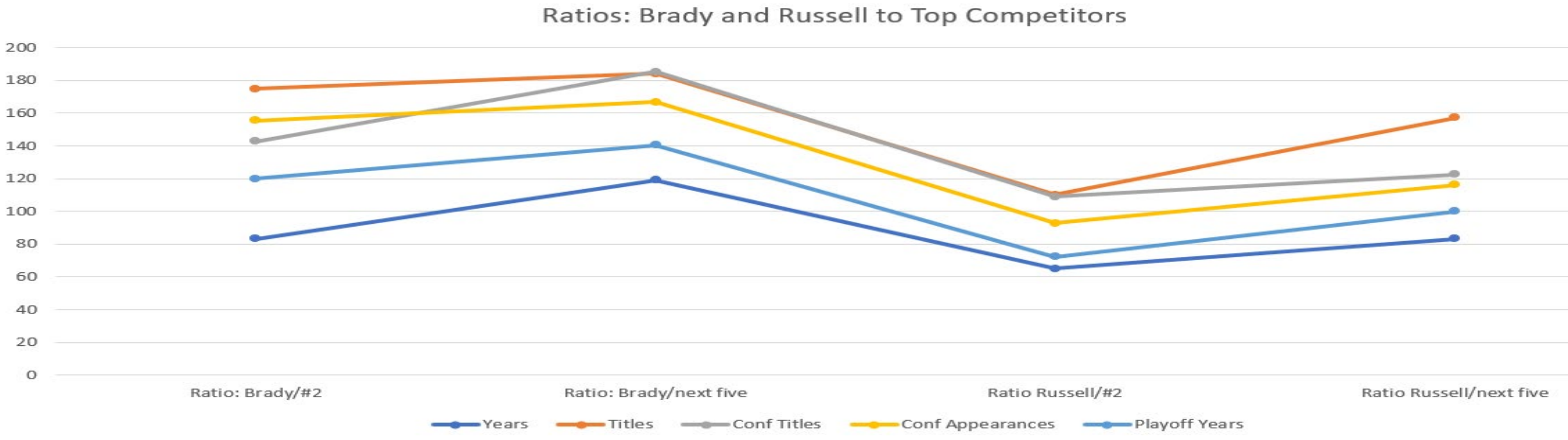
Vertical axis: P(stronger team wins series of 1,3,5,7 games) under independence.



BUT Brady played *GAMES;* Russell *SERIES*

Interpretation: Series help better teams avoid upsets. Also Brady had to win 3-4 games while Russell usually played two series. Does that make up the difference? IT MIGHT!

	Years	Titles	Championship	Conferences	Playoffs
Brady/#2*	83	175	143	156	120
Brady/next five	119	184	185	167	141
Russell/#2*	65	110	109	93	72
Russell/next five	83	157	122	116	100



Building Intuition: The $\log 5$ Method

Assume team A wins p percent of their games, and team B wins q percent of their games. Which formula do you think does a good job of predicting the probability that team A beats team B?

$$\frac{p+pq}{p+q+2pq}, \quad \frac{p+pq}{p+q-2pq}$$

$$\frac{p-pq}{p+q+2pq}, \quad \frac{p-pq}{p+q-2pq}$$

Building intuition: A wins p percent, B wins q percent

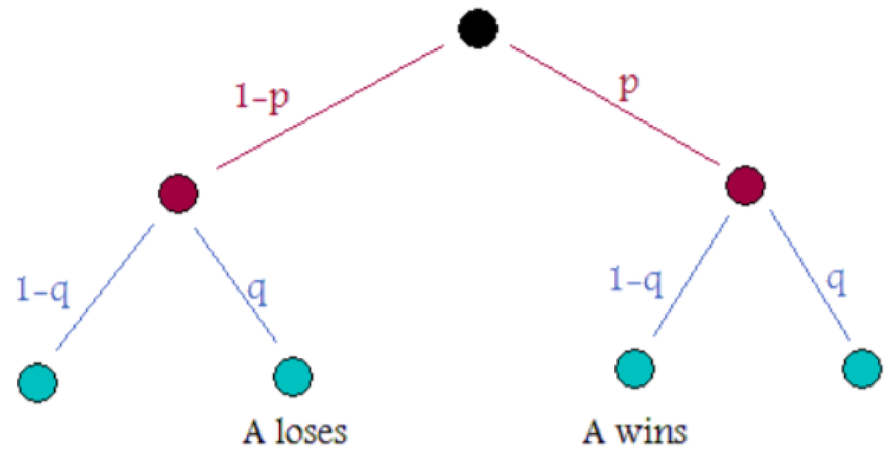
$$\frac{p+pq}{p+q+2pq}, \quad \frac{p+pq}{p+q-2pq}$$

$$\frac{p-pq}{p+q+2pq}, \quad \frac{p-pq}{p+q-2pq}$$

Consider special cases:

1. $\text{Prob}(A \text{ beats } B) + \text{Prob}(B \text{ beats } A) = 1.$
2. If $p = q$ then the probability A beats B is 50%.
3. If $p = 1$ and $q \neq 0, 1$ then A always beats B.
4. If $p = 0$ and $q \neq 0, 1$ then A always loses to B.
5. If $p > 1/2$ and $q < 1/2$ then $\text{Prob}(A \text{ beats } B) > p$.
6. If $q = 1/2$ prob A wins is p ($p = 1/2$ the prob B wins is q).

Building intuition: Sketch of proof: $\frac{p-pq}{p+q-2pq}$



- A beats B has probability $p(1-q)$.
- A and B do not have the same outcome has probability $p(1-q)+(1-p)q$.
- $\text{Prob}(A \text{ beats } B) = \frac{p(1-q)}{p(1-q)+(1-p)q} = \frac{p-pq}{p+q-2pq}$.

The log 5 rule (due to Bill James)

Suppose team A wins games with probability p, and team B wins with probability q.

A good estimate of P(A wins a game vs. B) is given by

$$\frac{p(1-q)}{p(1-q)+(1-p)q} = \frac{p - pq}{p+q - 2pq}$$

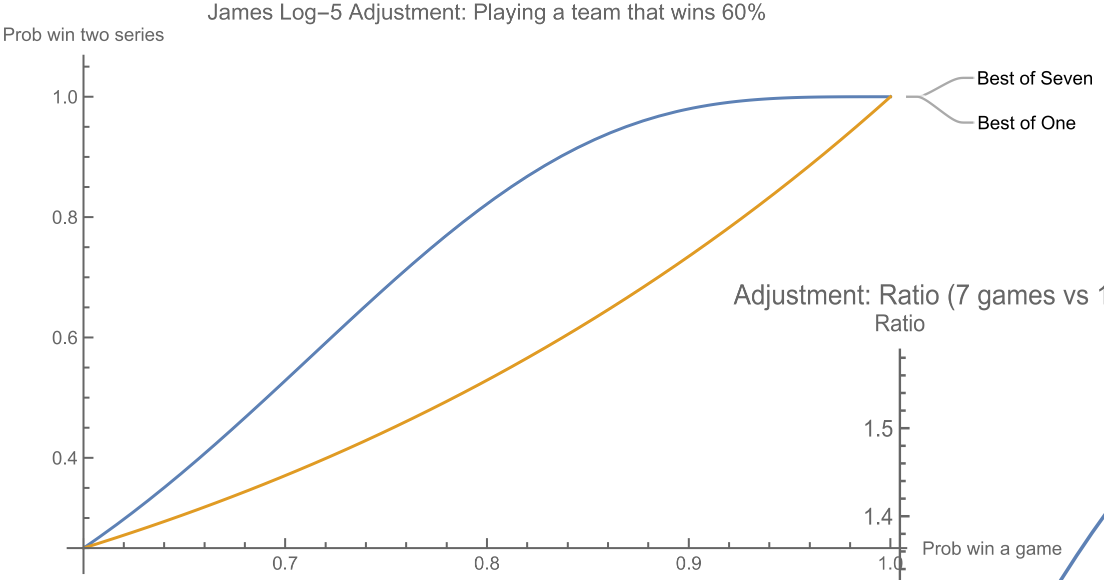
Example: In a playoff round we might have p = .8 and q = .6

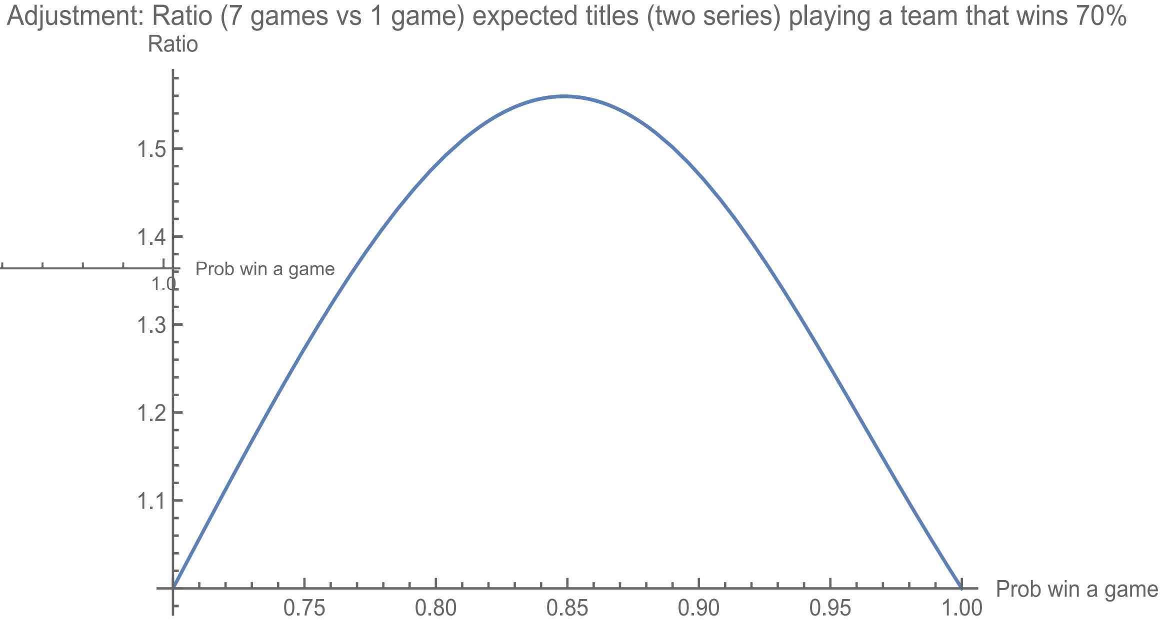
So P(A wins a game vs. B) = (.8 - .48) / (1.4 - .96) ≈ .727.

(See SJM's paper at
https://web.williams.edu/Mathematics/sjmiller/public_html/399/handouts/Log5WonLoss_Paper.pdf)

A Log-5 Adjustment for series vs. games

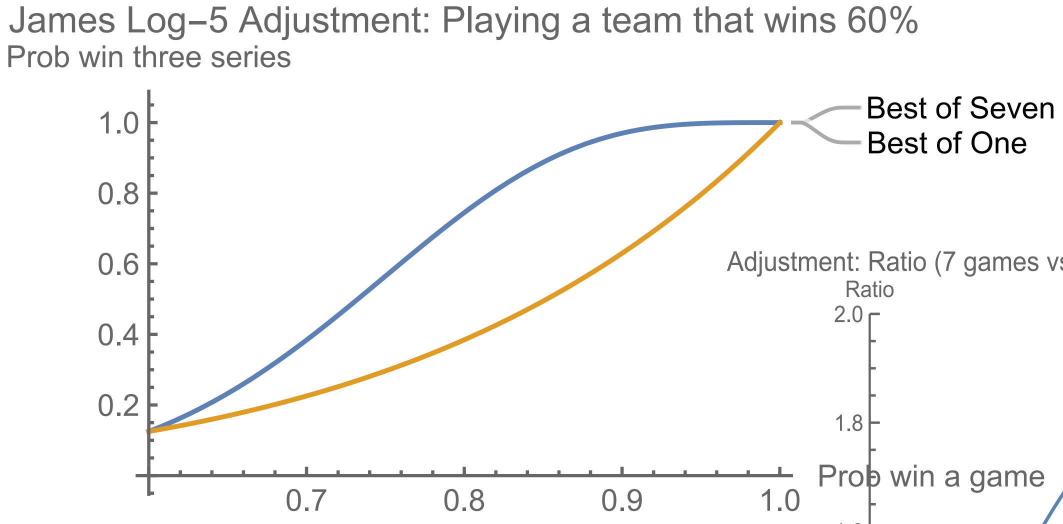
Team A, P(win)=p, plays B with q=0.6. Maximum ratio is about 1.6. Two series.

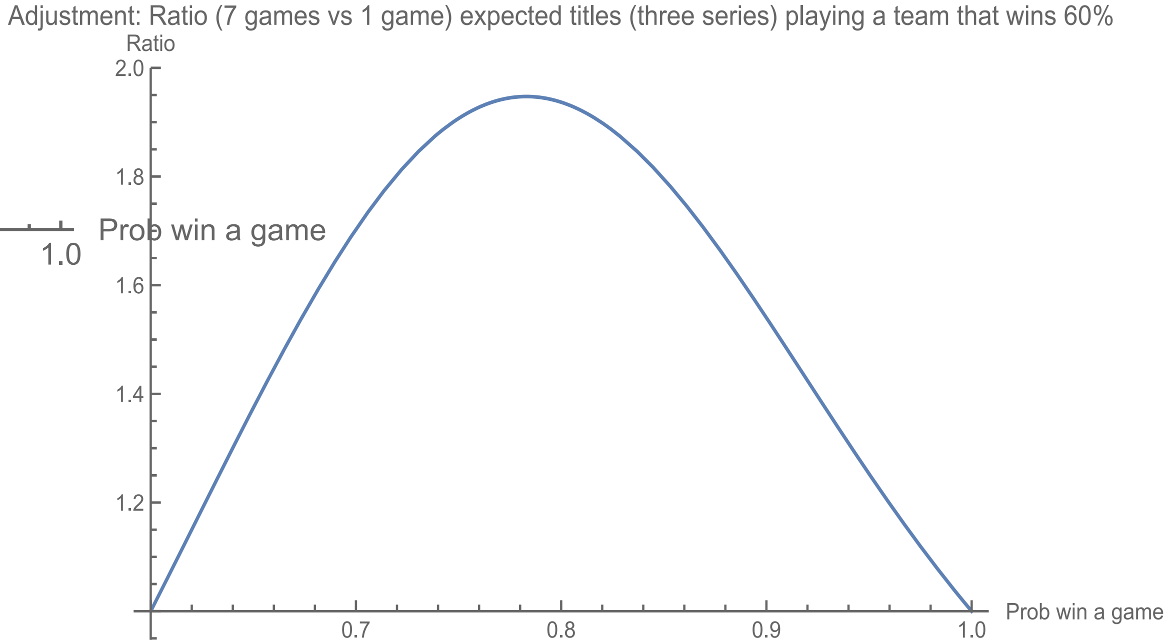




A Log-5 Adjustment for series vs. games

Team A, P(win)=p, plays B with q=0.6. Maximum ratio is about 1.9. Three series.





So here's an argument for Brady…

RUSSELL … 11 titles.

BUT NFL titles might be about 1.8 times more difficult because of the games/series issue!

So here's an argument for Brady…

RUSSELL … 11 titles.

BUT NFL titles might be about 1.8 times more difficult because of the games/series issue!

So BRADY's seven titles might be worth about:

$$7 * 1.8 \; = \; 12.6 \text{ RUSSELL titles!}$$

-Some next steps …

We have begun considering (and will recruit students to help us consider):

-More sophisticated metrics with playoff wins models (Poisson vs. binomial);

-Championship round success (proponents of Michael “six for six” Jordan and Joe “four for four” Montana

-Similar argument for longevity vs. high peak. Is four title in eight year career more or less impressive than four in 20?

-Influence by sport (Russell one of five, on floor 80% of time; Brady one of eleven, of field 40% of time … but Brady key in all of those plays while Russell might go some time without a touch.)

-We’ll think of more!

One more vote …

Based on our discussion, please use the chat to vote again for BOAT:

BRADY RUSSELL RUTH OTHER

Did we change any minds?
(We did change the question!)

THANKS for your attention ...

-Any questions please contact us at rcleary@babson.edu or sjm1@williams.edu

WHAT DO YOUMEAN?!?

Steven Miller, Williams College (sjm1@Williams.edu)

Definitions

Means and averages

- Given x and y, the average or mean is the number in between
- $\text{ArithmeticMean}(x,y) = (x + y) / 2$.
- There is more than one mean that can be defined!
- What properties should a mean have? Assume $0 < x \leq y$.

Desired Properties

We want:

- $x \leq \text{mean}(x,y) \leq y$. Should be "in between"
- $\text{mean}(x,x) = x$.

Does $\text{ArithmeticMean}(x,y) = (x+y)/2$ satisfy these properties?

Desired Properties

We want:

1. $x \le \text{mean}(x,y) \le y$. Should be "in between"
2. $\text{mean}(x,x) = x$.

Does $\text{ArithmeticMean}(x,y) = (x+y)/2$ satisfy these properties?

Proof of (1): Since $0 < x \le y$, we have $x + x \le x + y \le y + y$.
So we know $2x \le x + y \le 2y$. Divide everything by 2 and we get
$x \le (x+y)/2 \le y$ or $x \le \text{ArithmeticMean}(x,y) \le y$.
We proved the first result!

Desired Properties

We want:

1. $x \leq \text{mean}(x,y) \leq y$. Should be "in between"
2. $\text{mean}(x,x) = x$.

Does ArithmeticMean(x,y) = (x+y)/2 satisfy these properties?

Proof of (2):Does ArithmeticMean(x,x) equal x?

Yes! ArithmeticMean(x,x) = (x+x)/2 = 2x / 2 = x.

So the ArithmeticMean(x,y) = (x+y)/2 satisfies our two properties.

We write AM(x,y) = ArithmeticMean(x,y) = (x+y)/2 to save space.

Question

Is there another choice of mean that satisfies the two properties we wish?

We want:

1. $x \leq \text{mean}(x,y) \leq y$. Should be "in between"
2. $\text{mean}(x,x) = x$.

Thoughts?

Question

Is there another choice of mean that satisfies the two properties we wish?

We want:

1. $x \leq mean(x,y) \leq y$. Should be "in between"
2. $mean(x,x) = x$.

Try mean(x,y) = Sqrt(x y).

- Check: Sqrt(2 * 8) = Sqrt(16) = 4 and that IS between 2 and 8.
- Check: Sqrt(1 * 100) = Sqrt(100) = 10 and that is between 1 and 100.

So maybe this is another choice of mean. Maybe it also satisfies the two properties….

Question

Try mean(x,y) = Sqrt(x,y). Must show

1. $x \leq$ mean(x,y) $\leq y$. Should be "in between"
2. mean(x,x) = x.

First property: Show if $0 < x \leq y$ then $x \leq$ Sqrt(x y) $\leq y$.

We know $x \leq y$ so $x\,x \leq x\,y \leq y\,y$

But $x^2 \leq x\,y \leq y^2$. Now take the square-root!

Sqrt(x^2) = x and Sqrt(y^2) = y, so get $x \leq$ Sqrt(x y) $\leq y$, as claimed!

Question

Try mean(x,y) = Sqrt(x,y). Must show

1. x ≤ mean(x,y) ≤ y. Should be "in between"
2. mean(x,x) = x.

Second is easier!

We have Sqrt(x x) = Sqrt(x^2) = x. We are done!

We call this the GEOMETRIC MEAN. We write GM(x,y) = Sqrt(x y)

Two Means

So we have two choices of mean:

- AM(x, y) = (x + y) / 2
- GM(x, y) = Sqrt(x y)

BOTH have two good properties:

- For $0 < x \leq y$ both satisfy $x \leq \text{mean}(x,y) \leq y$ and $\text{mean}(x,x) = x$.

More used to the first.

Try x = 2 and y = 8:

- Get AM(2,8) = (2 + 8) / 2 = 10 / 2 = 5
- Get GM(2,8) = Sqrt(2 * 8) = Sqrt(16) = 4

Two Means

So we have two choices of mean:

- AM(x, y) = (x + y) / 2
- GM(x, y) = Sqrt(x y)

BOTH have two good properties:

- For $0 < x \leq y$ both satisfy $x \leq \text{mean}(x,y) \leq y$ and $\text{mean}(x,x) = x$.

More used to the first.

Try x = 3 and y = 12

- Then AM(3, 12) = 15/2 = 7.5
- And GM(3,12) = Sqrt(36) = 6.

Two Means

So we have two choices of mean:

- AM(x, y) = (x + y) / 2
- GM(x, y) = Sqrt(x y)

BOTH have two good properties:

- For $0 < x \leq y$ both satisfy $x \leq \text{mean}(x,y) \leq y$ and mean(x,x) = x.

Try x = 1 and y is VERY large....

- Then AM(1,y) = (1 + y)/2 which is APPROXIMATELY y/2
- But GM(1,y) = Sqrt(y) which is MUCH smaller if y is large.
- Note if y is small we would say (1 + y)/2 is approximately .5

CONJECTURE: GM(x,y) ??? AM(x,y)

Two Means

So we have two choices of mean:

- AM(x, y) = (x + y) / 2
- GM(x, y) = Sqrt(x y)

BOTH have two good properties:

- For $0 < x \le y$ both satisfy $x \le \text{mean}(x,y) \le y$ and mean(x,x) = x.

Try x = 1 and y is VERY large….

- Then AM(1,y) = (1 + y)/2 which is APPROXIMATELY y/2
- But GM(1,y) = Sqrt(y) which is MUCH smaller if y is large.
- Note if y is small we would say (1 + y)/2 is approximately .5

CONJECTURE: $GM(x,y) \le AM(x,y)$

CONJECTURE: GM(x,y) ≤ AM(x,y)

PROOF: Consider: $0 < x \leq y$, what is true about $(\text{Sqrt}(x) - \text{Sqrt}(y))^2$? It must be positive…

- So $0 \leq (\text{Sqrt}(x) - \text{Sqrt}(y))^2$.

Remember FOIL: $(a - b)^2 = (a - b)(a - b) = a a - a b - b a + b b$: First Outside Inside Last

- So $(a-b)^2 = a^2 - 2 a b + b^2$

We are looking at $(\text{Sqrt}(x) - \text{Sqrt}(y))^2$.

- $0 \leq (\text{Sqrt}(x) - \text{Sqrt}(y))^2 = \text{Sqrt}(x)^2 - 2\, \text{Sqrt}(x)\, \text{Sqrt}(y) + \text{Sqrt}(y)^2$.
- $0 \leq x - 2\, \text{Sqrt}(x\, y) + y$

Trying to get $AM(x,y) = (x+y)/2$ and $GM(x,y) = \text{Sqrt}(x,y)$

- $2\, \text{Sqrt}(x,y) \leq x + y$
- $\text{Sqrt}(x,y) \leq (x+y)/2$
- $GM(x,y) \leq AM(x,y)$.

We proved it!

Extensions

What if we had three objects: $0 < x \leq y \leq z$?

- $AM(x,y,z) = (x+y+z) / 3$
- $GM(x,y,z) = (x\ y\ z)^{1/3}$.

Is there another combination?

- $((x\,y\ +\ y\,z\ +\ x\,z) / ???\ )^{???}$

Food for thought: can you find a choice of a and b such that

- $((xy + yz + zx) / a)^b$ is a mean, so it would satisfy
- $x \leq TripleMean(x,y,z) \leq z$ and $TripleMean(x,x,x) = x$

If $x = y = z$ then $((xx + xx + xx) / a)^b = (3\ x^2 / a)^b = x$ for ALL x.

- SO b = ??? and a = ???

Extensions

What if we had three objects: $0 < x \leq y \leq z$?

- $AM(x,y,z) = (x+y+z) / 3$
- $GM(x,y,z) = (x\ y\ z)^{1/3}$.

Is there another combination? YES

- $((x\ y\ +\ y\ z\ +\ x\ z) / 3\ )^{1/2}$

If $x = y = z$ then $((xx + xx + xx) / a)^b = (3\ x^2 / a)^b = x$ for ALL x.

- SO $b = ½$ and $a = 3$

SO this is our guess....

- Try $x = 3$ and $y = 4$ and $z = 5$
- $TripleMean(3,4,5) = (\ (12 + 20 + 15) / 3\ )^{1/2} = (47/3)^{1/2}$ is approximately 3.958
- This IS a reasonable answer! It is more than 3, less than 5!

Final Thoughts

AM(x,y) = (x+y)/2   GM(x,y) = Sqrt(x y)

Test 1 Get 1 and on Test 2 get 100

- AM(1, 100) = (1 + 100)/2 = 50.5
- GM(1,100) = Sqrt(1 100) = 10

Recall

- Log(x y) = Log(x) + Log(y)
- So there is a relation between logarithms, AM and GM

https://kconrad.math.uconn.edu/articles/maclaurin.pdf

Maclaurin's Inequality and a Generalized Bernoulli Inequality Iddo Ben-Ari &Keith Conrad

Slides below here are for another day…

Players to consider… NFL

		Yrs	Titles	FInals	Semis	Playoffs
Tom Brady	2002-21	20	7	10	14	18
Joe Montana	1980-94	15	4	4	8	11
Terry Bradshaw	1971-84	14	4	4	6	9
Jerry Rice	1986-05	20	4	4	8	14
Steve Young**	1986-00	15	3	7	7	11
Roger Staubach	1970-80	11	2	4	7	10
John Elway	1984-99	16	2	5	6	10
Adam Vinatieri	1997-20	24	4	5	6	15

Players to consider... NBA

	Seasons	Yrs	Titles	Final	Semis	Playoffs
Bill Russell	1958-70	13	11	12	13	13
Michael Jordan	Various	15	6	6	8	13
Sam Jones	1958-69	12	10	11	12	12
K Abdul-Jabbar	1971-90	20	6	10	14	18
LeBron James	2004-21	18	4	10	11	9
Tom Heinsohn	1957-65	9	8	9	9	9
Magic Johnson	Various	13	5	9	10	13

GOAT METRICS IN TEAM SPORTS …

- Individual statistics?
  - -Counting stats like Points, Touchdown Passes, Home Runs, Goals, Assists
  - -Derived stats like Wins Above Replacement, Adjusted Plus/Minus
- Team statistics?
  - -Titles won (consider time periods, league size)
  - -Consistent excellence (high winning percentage over a long time.)

Measuring GOATS by their tails...

Principle: What a GOAT does is accomplish something that is *least likely* among a set of candidates using a particular metric. There are many variables to consider when computing these probabilities, but as an exercise...

Positives:

-Great modeling project that requires careful consideration of inputs and assumptions.

-Adaptable; can be used for both team and individual performance.

Measuring GOATS by their tails…

Principle: What a GOAT does is accomplish something that is *least likely* among a set of candidates. There are many variables to consider when computing these probabilities, but as an exercise…

Positives:

-Great modeling project that requires careful consideration of inputs and assumptions.

-Adaptable; can be used for both team and individual performance.

Negatives:

-Sort of like a p-value. Why right tail only?

-Almost certainly not definitive.

The “What Abouts…”

- What about teammates? Russell played 80% of the time and was one fifth of the team at any time. Brady played 40% of all plays and was 1/11 of his team at any time.

Another one we might not have time to explore in talk

The “What Abouts…”

- What about influence? Russell played 80% of the time and was one fifth of the team at any time. Brady played 40% of all plays and was 1/11 of his team at any time.
- What about league parity? The NBA in Russell’s era had teams that were never truly competitive.

The "What Abouts…"

- What about influence? Russell played 80% of the time and was one fifth of the team at any time. Brady played 40% of all plays and was 1/11 of his team at any time.
- What about league parity? The NBA in Russell's era had teams that were never truly competitive.
- What about quality of teammates? (And many more…)

So stay tuned: We hope to attract students to build these more sophisticated interpretations!