Top 8th

The Numbers Game: What Fans Should Know About the Stats They Love

Jay Bennett and Aryn Martin

There is an old *Peanuts* cartoon in which Charlie Brown, the unluckiest player/manager in baseball, gives up a walk-off home run in the last game of the season. His hapless team ends the year with 0 wins and 20 losses. Some time later, Schroeder, the team's statistician, consoles his sad-sack pitcher.

"I've worked up some interesting statistics here about our baseball team Charlie Brown…

"I think you'll find that they say something to us…

"Last year our opponents scored three thousand and forty runs to our six! They made forty-nine hundred hits to our eleven and they made nineteen errors to our three hundred…"

Charlie Brown, disheartened, finally shouts, "Tell your statistics to shut up!!"[1]

Playing the Numbers Game

Baseball fans (with the possible exception of Charlie Brown) love statistics. More than in any other sport, fans of America's pastime pore over tables to venerate and vilify players, teams, and managers. Numbers seem to transcend time and place as they allow fans to compare events and players that are miles and years apart. Why are baseball fans so fond of stats? Is it simply because the numbers tell the story about streaks and slumps? Is it because fans are drawn into the game as participants and amateur mathematicians? Is it because, during a slow game, scribbling on the scorecard passes the time between trips to the hot dog stand? "Baseball fans are junkies, and their heroin

[1] Charles M. Schulz, *Slide, Charlie Brown! Slide!* (Greenwich, Conn.: Fawcett Pulications Inc., 1962).

is the statistic."[2] You can miss a game and still get a quick fix of ERAs, RBIs and HRs to prevent withdrawal symptoms. In the stands, our expectations rise and fall with the array of statistics on the digital scoreboard, as though they tell us something very tangible and important about a player's past, present, and future.

Baseball fans have some things in common with another community of avid counters and enumerators. All manner of scientists are daily involved in the generation of numerical constants to describe or predict properties of nature. However, as the philosopher and historian of science Thomas Kuhn (1922-1996) suggests, this has not always been the case.[3] He describes a second scientific revolution (the first having occurred in the 17th century with Bacon, Galileo, Newton et al), in which quantification was a central feature. During the mid 1800s, Kuhn argues, physical scientists became newly concerned with generating an "avalanche of numbers." Philosopher of science Ian Hacking summarizes Kuhn's argument: "The world was now conceived in a more quantitative way than ever before. The world is seen as constituted by numerical magnitudes."[4] As an example, Hacking cites a pamphlet published in 1835 by Charles Babbage, the father of computing. In it, Babbage urged the publication of tables of all the numerical constants known in the sciences and the arts, from specific gravities and atomic weights to the amount of oak a man can saw in an hour. The everyday practice of science in the mid to late 1800s--and arguably today--was aimed at devising precise and

[2] Robert S. Weider, quoted on http://www.baseball-almanac.com/quotes/stats5.shtml.
[3] Thomas S. Kuhn, "The Function of Measurement in Modern Physical Science," *Isis,* Vol. 52, Issue 2, 1961, pp. 161-193.
[4] Ian Hacking, *Representing and Intervening: Introductory Topics in the Philosophy of Natural Science* (Cambridge: Cambridge University Press, 1983), p. 242.

ingenious instruments and methods for "obtaining very accurate numbers that don't matter much."[5]

Perhaps it is not coincidental that baseball, a sport afloat in statistics, originated in the mid 1800s, when the scientific world was engaged in a proliferation of measurement. If the urge to enumerate in baseball is similar to what Hacking describes as "the fetish for measuring precise numbers" in the sciences, maybe the motivations are also analogous. Kuhn suggests that with the explosion of statistics in physics came a decrease in the length of controversies about scientific theories, and an increase in consensus that emerged from such controversies. Imagine a debate about the best hitter in baseball without numbers. On what would one base an argument but personal proclivities and scattered observations? Bringing batting averages into the debate shortens controversy and promotes consensus. For example, to answer the question *who was the best hitter in the National League in 2003?*, all we have to do is consult last year's voluminous statistics and we find that Barry Bonds won the batting championship with the highest batting average of .370. It's all very simple. Just sorting a set of numbers.

What could be complex about statistics, and in particular, baseball statistics? After all, statistics are just numbers and numbers are a definitive means of answering a question. But for statisticians and philosophers, these numbers are complex elements in themselves. They are a synthesis and view of the complex process that created them, an intricate sequence of balls, strikes, line shots, groundballs, running catches, and sprints to the next base on the path home. Vast compilations of batting and pitching statistics in encyclopedias and Internet websites give false comfort in the precision of how well we

[5] Ibid., p. 236.

understand player and team abilities. One of the ironies of statistical analysis is that it often reveals the limitations of our knowledge.

At the end of each season, Major League Baseball (MLB) recognizes the player with the highest batting average in each league as the league's batting champion. The batting average (BA) is simply the number of hits (H) divided by the number of at bats (AB). Over the course of MLB history, there have been many tight races for this championship. As recently as 1998, Mo Vaughn (.337 BA) lost the AL batting championship to Bernie Williams (.339 BA). In the last 50 years, six out of the 100 batting championships have been decided by the result of one at bat. That is, if the challenger had just one more hit, he would have displaced the winner as batting champion. The table below lists these seven close races.

Year	League	Player	Team	AB	H	BA	Difference
2003	National	Albert Pujols	St. Louis	591	212	0.3587	0.0002
		Todd Helton	Colorado	583	209	0.3585	
1970	American	Alex Johnson	California	614	202	0.3290	0.0004
		Carl Yastrzemski	Boston	566	186	0.3286	
1982	American	Willie Wilson	Kansas City	585	194	0.3316	0.0009
		Robin Yount	Milwaukee	635	210	0.3307	
1991	National	Terry Pendleton	Atlanta	586	187	0.3191	0.0011
		Hal Morris	Cincinnati	478	152	0.3180	
2003	American	Bill Mueller	Boston	524	171	0.3263	0.0012
		Manny Ramirez	Boston	569	185	0.3251	
1976	American	George Brett	Kansas City	645	215	0.3333	0.0013
		Hal McRae	Kansas City	527	175	0.3321	
1960	National	Dick Groat	Pittsburgh	573	186	0.3246	0.0019
		Norm Larker	Los Angeles	440	142	0.3227	

For example, in 1991, if Hal Morris had one more hit in his 478 at bats, he would have finished with a .320 BA, ahead of Terry Pendleton. When you consider the number of close plays that could occur over the course of a season, isn't it possible that one of his 326 outs could easily have been a hit? A bang-bang play at first that didn't go his way?

Or a line shot barely speared by a leaping infielder? Perhaps it was pure chance that Morris did not get that extra hit to win the batting championship. On the other hand, maybe Pendleton could have had one less hit which would have given him a .317 BA. Perhaps he was fortunate in a ruling that gave him a hit where an error would have been in order.

Morris was not Pendleton's only serious challenger for the 1991 NL batting championship. San Diego outfielder Tony Gwynn finished with a .317 BA. Two additional hits in his 530 at bats would have given him the title. Maybe he actually was a better hitter than Pendleton but luck was not in his favor. From a statistician's perspective, there is an element of chance in any measurement (such as batting average) resulting from a process (such as a baseball season). For a statistician, luck/chance (the terms will be used synonymously) is what is left over after the elements that are controlled or understood are eliminated.

The Luck of the Average

It is not difficult to imagine that luck or chance may have been the final arbiter in determining which of these three players won the 1991 NL batting championship. The same goes for the other six closely contested races. But statistical theory indicates that the set of potential contenders may be much larger.

The theory is based on the premise that each batter has an *ability* to obtain a hit in each at bat. This ability has an unknown constant numerical value. The outcome of each at bat is a random event--that is, one determined by chance subject to constraints from the batter's ability and other possible factors (such as the ability of the opposing pitcher, characteristics of the ballpark, and the vagaries of the weather). For simplicity, we will focus on the batter's ability by assuming that these other factors are either not significant or folded into the chance element.

Under this theory, if a batter's ability is marked at .300, he has a 30% chance to get a hit and a 70% chance of making an out in each at bat. This is the easy part of the theory. The difficult part of the theory arises from the "unknown" aspect of the definition of ability. The standard assumption is that if a batter gets 30 hits in 100 at bats, then his ability is .300. However, based on statistical theory, in 100 at bats, a batter with .250 ability has a 5% chance of getting 30 hits, a 3% chance of getting 31 hits, and even a very small chance of getting 100 hits. If we consider all possibilities of getting 30 to 100 hits, he has a 15% chance of getting 30 hits or more. So, if we observe a batter with 30 hits in 100 at bats, it is very likely that his ability is close to .300 but there is still a substantial possibility that it could be .250 or even lower.

Of course, over an entire season, regular position players have many more than 100 at bats. According to the *Official Rules of Major League Baseball*, a player in a 162-game season must have at least 502 plate appearances to qualify for the batting championship. Plate appearances include official at bats, along with walks, hit by pitches, sacrifice flies, sacrifice hits, and catcher's interference. (Barry Bonds only had 403 official at bats in 2002, but his 198 walks elevated his plate appearances to qualify for the National League batting championship.) This greatly increased sample size might be expected to have a powerful effect in reducing the significance of chance in batting averages of regular players.

Continuing with our 2002 Bonds example, we observed that Bonds had 149 hits in 403 at bats for a .370 BA. However, according to statistical theory, while .370 is our best estimate of his batting ability, there is a 16% chance that his ability was actually lower than .346 and only good luck gave him a .370 BA. It works both ways so that there is also a 16% chance that his ability was higher than .394 and bad luck reduced his BA to .370. It is important to distinguish between his ability (which is fixed but unknown) with his BA (which is known from observation but is at least partly the result of chance). Generally, the two are treated as being identical, but they are not. The observed BA is only a clue (albeit an important one) to finding his true ability.

Larry Walker was runner-up to Bonds in the 2002 NL batting championship. With a .338 BA, Walker finished a distant 32 points behind Bonds. Could luck have accounted for such a large difference in BA? Actually, statistical theory indicates that there is about a 16% chance that Walker was the better hitter in 2002 and only Barry's good luck and Larry's bad luck made the difference. (Recall that we are using a simple model here that does not account for park effects which could be substantial in any comparison with Walker who played for Colorado.) If luck could have been significant, imagine what that implies about the six close races discussed earlier. In effect, those 6 races were decided by luck; the first and second place batters had virtually the same batting abilities.

How powerful is the influence of luck in baseball? Is it possible that *all* Major League hitters have the same ability and only chance determines their final positions in the race for the batting championship? In 2002, 146 players in the American and National Leagues qualified for the batting championship with 502 plate appearances.

Their BAs ranged from .215 (Jeromy Burnitz, Mets) to .370 (Bonds) with .278 being the average value. Since we are only considering hitters with a substantial number of plate appearances, marginal players are not included. Therefore, the hypothesis that all have the same ability is not as preposterous as it might seem. Theoretically, if all 146 batters had the same ability (.278 BA), we would expect about 18 hitters to bat higher than .300 and about 10 to bat less than .250 in a season by chance alone.[6] In 2002, 32 hitters had a BA greater than .300 and 25 had a BA less than .250. Since many more batters were observed in the extreme ends of BAs than expected from chance alone, the distribution of observed BAs has greater spread than can be accounted from chance. It is reasonable to conclude that differences in batting ability play a major role in determining the batting champion as well as chance.

We can actually get a rough idea of the degree to which ability and luck play a role in determining a player's batting average. The variation in observed BAs from .215 to .370 is the result of two sources of variation: ability and chance. We calculated the variation due to chance so whatever is leftover is due to ability. This is a reverse application of Branch Rickey's oft-quoted axiom that "Luck is the residue of design."[7] We are removing the luck factor and assuming that what is left is variation in ability from player to player. Since the variation of BAs due to chance is about half that observed in the 146 player BAs, we conclude that chance and ability played equal parts in determining player BAs observed in 2002. Eliminating chance from the observed BAs,

---

[6] This calculation was performed using a simulation of 1000 batters all with the same ability in getting hits. In this simulation the number of at bats was randomized from the set of at bats for the 146 players considered. This simulation indicated that if all 146 players had the same ability the mean batting average would be .278 and the standard deviation of batting averages would be .01909. Based on a gamma distribution with this mean and standard deviation, 18 of the 146 players would have observed batting averages greater than .300 and 10 players would have observed batting averages less than .250.

[7] Paul Dickson, *Baseball's Greatest Quotations* (New York: HarperCollins, 1991), p. 356.

we can estimate that the true abilities of the 146 players are most likely in the range from .230 to .331, rather than .215 to .370.

Batting average is the oldest and most familiar measure of batting performance in baseball. However, many studies have demonstrated that BA is perhaps the weakest of all batting measures in estimating run production.[8] The on-base percentage (number of times reaching base divided by the number of plate appearances) and the slugging percentage (total bases from hits divided by at bats) improve on BA in this respect. Summing these two values together to obtain on-base plus slugging (OPS) is better than either separately. Does luck have the same influence over other batting measures?

If we perform a similar analysis on on-base percentage, slugging percentage, and OPS as we did for BA using the 146 qualifiers for the batting championships in 2002, we find that ability played a much larger role in determining their observed values, about three to four times that of chance. This evidence provides further support for their use instead of BA. Not only are they better measures of run production, but they are also better in discriminating ability from chance.

Great Teams or Just Plain Lucky?

Individual player performances are of great interest to fans, but the heart of any sport lies in the team performance, particularly in their quest for a championship. How much does chance affect baseball championships?

Throughout its history, baseball has had its share of miracle teams who surprised everyone, fans and sportswriters alike, in their great achievements beyond all

[8] See for example John Thorn and Pete Palmer, *The Hidden Game of Baseball* (Garden City, New York: Doubleday, 1985) and Jim Albert and Jay Bennett, *Curve Ball: Baseball, Statistics, and the Role of Chance in the Game* (New York: Springer-Verlag, 2001).

expectations. Perhaps the most famous such team in recent history is the 1969 New York Mets, who won 100 games (most in the National League), swept the Atlanta Braves in the League Championship Series, and won the World Series from the heavily favored Baltimore Orioles in five games. From their record, the Mets sound like an indomitable team. But it didn't appear that way when the 1969 season opened. The 1968 Mets finished one game out of the cellar in a ten-team National League; their .451 winning percentage was their best finish to date in their brief seven-year history. In 1970, they returned to mediocrity with an 83-79 record. Were the Mets truly the best team in baseball in 1969, or were they just lucky?

For the book *Curve Ball*, Jim Albert developed a statistical model of team performance based on historical distributions of team talent. He designed an experiment in which 1000 baseball seasons were replayed using the current structure of 162-game seasons, followed by a playoff tournament involving six divisional champions and two wild card teams. In each experimental season, each team is randomly assigned a talent within the constraints of historical distributions of team talent. In each game, a victor is determined by comparing the talents of the two teams subject to an element of chance. Unlike the historical MLB record where only the performance is known, both performance and talent are known for the results of these 1000 virtual seasons. So, we can see how well a team with a certain level of talent can be expected to perform. A team with average talent should only win about half of its games. And yet Albert's experiment indicates that such a team still has a 19% chance of getting into the playoffs, a 3% chance of winning the pennant, and a 1% chance of winning the World Series. Even more startling is that the team with the best talent of all 30 teams has an 11% chance of missing the playoffs and only a 21% chance of winning the World Series. According to Albert's simulation, about 4 of every 5 World Series under the current season and playoff structure is *not* won by the most talented team!

We can also see the converse: how much performance can tell us about a team's talent. If a team wins the World Series, how good was the team apart from its luck? Almost half of World Series champions are among the top three talent-laden teams in a MLB season. Surprisingly though, more than one out of eight World Series champions has only average talent or worse. ESPN columnist Rob Neyer observed that "winning the World Series isn't about being the best, it's about being the luckiest."[9]

Much of this is a result of the playoff structure. The existence of a wild card team was introduced as a safety net allowing a strong team less favored by its divisional placement to have an opportunity in the playoffs. Albert's results indicate that this system rescues the team with best talent once every eight seasons. On the other hand, for every such great team given a second chance, almost five teams with average or worse talent are given an opportunity that they may not deserve. Even without the inclusion of a wild card team, the divisional structure gives less talented teams greater opportunity to enter the playoffs where the limited number of games allows chance a greater role. The distribution of talent across divisions is just another way that chance is introduced into the determination of a World Series champion.

Thus, the current MLB structure may not be optimal for determining which team is best. One might ask if the purpose of playing MLB games is to find the best team or is it to entertain the MLB fan base? Clearly, the games are played as entertainment. The divisional organization and playoff structure do provide a greater opportunity for less

---

[9] http//www.espn.com, April 2, 2003.

talented teams to overcome more talented ones. Whether this makes baseball more entertaining is open to conjecture and may depend on personal preference. The current structure of playoffs guarantees a climatic series of games at the conclusion of the season at the expense of never having the season long suspense of a pennant race for all the marbles in a league. We will have great championship series like the '80 NLCS in which the Phillies topped the Astros in five games, four of which went into extra innings including the finale. But we will never again have the '51 NL pennant race in which the New York Giants overcame a mid-August 13 1/2 game deficit to the Brooklyn Dodgers to tie them at the end of the regular season and won a three-game playoff with Bobby Thompson's "shot heard around the world" on the final at bat.

In general, the element of chance, the unexpected, does appear to add to the entertainment value of baseball. However, too much chance, the appearance that the "best" team does not win often enough devalues the games. Some compromise between the two is best. The point at which that compromise lies is still an open issue.

Numbers and Players

Perhaps players have an intuitive understanding that the luck of the numbers plays such a big part in their professional accomplishments. Baseball players are noted for being the most superstitious of professional athletes. This perceived lack of ability to control their destiny may be at the root of the many superstitions that permeate baseball. Mets and Phillies outfielder Lenny Dykstra was noted for throwing out his batting gloves and changing his wad of chewing tobacco if he made an out.[10] This perception may well be

---

[10] Robert Gordon and Tom Burgoyne, *More Than Beards, Bellies, and Biceps: The Story of the 1993 Phillies (And the Phillie Phanatic Too)* (Champaign, IL: Sports Publishing L.L.C., 2002), p. 62.

influenced by the limited control they feel in the outcome of a play. They have confidence in their abilities, but they recognize that their skills may take them only so far. Hall of Fame pitcher Lefty Gomez claimed he would rather be lucky than good.[11] Statistics show that while Gomez may be guilty of hyperbole, there is a solid rationale for his conviction.

Continuing in this vein, managers have less control over game outcomes than players, and as a group are reputed to be even more superstitious. Sparky Anderson never stepped on a foul line when visiting the mound. Gene Mauch never had his uniform cleaned when his team won.[12] (At least the ‘64 Phillies, losers of the National League pennant when they dropped 10 of their last 12 games, sank with a skipper in a clean uniform.) In 1911, Charles Victory Faust told John McGraw that a fortune teller had guaranteed the New York Giants would win the pennant if he pitched for them. Although Faust had no skill whatever as a pitcher, McGraw kept him on the Giant payroll from 1911 through 1913 as a good luck charm. Faust warmed up for every game (though he never started) and the Giants did win the pennant in each of those years.[13]

McGraw, who was nicknamed Little Napoleon, may have been inspired by Bonaparte himself who was reputed to have implored “Give me generals who are lucky!” So, the notion that luck potentially plays a major role in success cuts across all strata and professions. What makes baseball so interesting is that the records kept of performance and the situations in which they occur are more extensive and thorough than in any other profession. They allow us to truly get a quantitative sense of the extent to which baseball player careers (and our own lives) are affected by chance.

---

[11] Bob Chieger, *Voices of Baseball* (New York: New American Library, 1983), p. 153.
[12] http://www.oaklandchamber.com/html/2000_04_As_superstition.html.
[13] Lawrence S. Ritter, *The Glory of Their Times* (New York: Macmillan, 1966), pp. 93-97.

Numbers as Players

Numbers, it is oftentimes supposed, strip away subjective opinion in favor of cold hard facts. They turn a chaotic world into orderly tables expressed in a universal and crisp language. They lend an air of irrefutable authority to any claim. As Lord Kelvin famously remarked, "When you measure what you are speaking about and express it in numbers, you know something about it; but when you cannot measure it in numbers, your knowledge is of a meager and unsatisfactory kind.[14] Numbers, then, are appealing to baseball fans and scientists alike because they seem to crystallize information so that it can be readily communicated, compared and exchanged.

But it's easy to forget that numbers are wily. Philosophers of science can also be instructive in helping us to ponder some less than obvious characteristics of letting the numbers tell the story. At first glance, baseball statistics, like specific gravities and atomic weights, seem to be purely descriptive: they describe elements of the world in bland and no-nonsense ways. If numbers are descriptive--if statistics reflect what happened in a baseball game--they are at best partially so. Porter reminds us that:

> Mathematical and quantitative reasoning . . . provide no panacea. Mapping the mathematics onto the world is always difficult and problematical. Critics of quantification . . . have often felt that reliance on numbers simply evades the deep and important issues.[15]

Numbers embody all manner of contingencies and subjective moments that lie buried in their making. In particular, the role of chance in producing these numbers (as demonstrated with batting averages) is neglected. The size of the ball park and the

---

[14] Quoted in Theodore Porter, *Trust in Numbers* (Princeton: Princeton University Press, 1995), p.72.
[15] Ibid., p. 5.

particular pitchers faced will impact a hitter's average. Also, factors such as the wind and shadows are not in the hitters' control (though a seasoned hitter will make an effort to adjust). Even human decisions, such as error scorings can affect the way a hit is recorded and thus become part of the batting average process. As we saw earlier, when statistics are calculated differently--on-base plus slugging instead of batting average, for example--they subtly change how events are described and understood. As Hacking asks, "Do measurements measure anything real in nature, or are they chiefly an artifact of the way in which we theorize?"[16] Statistics don't lie, but they are subject to misinterpretation. A statistical description of a baseball game may be like "a travel book that ignored a charming landscape and its inhabitants in favor of recording precisely the times of arrival and departure of trains."[17]

Some philosophers would propose that numbers are not descriptive but prescriptive: they shape the world they are meant to describe. An anthropologist can change an environment simply by observation. Even in physics, as Neils Bohr (1885-1962) famously observed, measurement itself becomes a part of the phenomenon being measured.[18] Similarly, Porter suggests that "measures succeed by giving direction to the very activities that are being measured."[19] In other words, numbers have agency in the world, they do things. Porter draws on the work of French philosopher Michel Foucault (1926-1984) to assert that "numbers have often been an agency for acting on people, exercising power over them. . . . Numbers turn people into objects to be manipulated.

---

[16] *Representing and Intervening*, p. 233.
[17] *Trust in Numbers*, p. 18. Porter refers to this comparison made in the 1830s by the Hegelian natural philosopher Georg Friedrich Pohl to describe Georg Simon Ohm's mathematical treatment of the electrical circuit.
[18] In his elaboration of the Heisenberg Uncertainty Principle, Neils Bohr observes that "measurement has an essential influence on the conditions on which the very definition of the physical quantities in question rests." See his "Quantum Mechanics and Physical Reality," in *Nature*, Vol. 136, 1935, p. 1025.
[19] *Trust in Numbers*, p. 45.

Where power is not exercised blatantly, it acts instead secretly, insidiously."[20] We can see this in baseball where statistics determine decisions such as player drafting and trading, salaries, the hiring and firing of managers, not to mention Las Vegas odds. Milestone numbers are particularly important in shaping the game; among other things, they strongly influence (not without controversy) entrance into the Hall of Fame. Ticket sales soared when Roger Clemens pitched for his 300th victory, although the game itself may be no better than his 299th or 301st. Consider also the magical 500 home run mark for admittance into the Hall of Fame. Statistics have the capability to illuminate the mysteries of the game and act as a force for improvement in the sport. But, despite their seeming descriptive innocence, their misuse can obscure our understanding and orchestrate elements to the detriment of the game.

So, Charlie Brown you would be wise to heed Schroeder's statistical analysis. But take heart in the success of the 1914 Braves, the 1969 Mets, and the 2002 Angels. Wait 'til next year and don't step on the foul lines!

---

[20] Ibid., p. 77.