

Journal of Quantitative Analysis in Sports

Volume 2, Issue 1

2006

Article 2

Pitching Statistics, Talent and Luck, and the Best Strikeout Seasons of All-Time

James Albert*

*Bowling Green University, albert@math.bgsu.edu

Copyright ©2006 The Berkeley Electronic Press. All rights reserved.

Pitching Statistics, Talent and Luck, and the Best Strikeout Seasons of All-Time

James Albert

Abstract

Many count statistics are used to evaluate pitchers such as the number of wins and losses, the number of strikeouts, the number of walks, and the number of runs allowed. For a given measure such as strikeouts, this paper focuses on the estimation of pitchers' probabilities of striking out a batter. The variation in the season strikeout rates among a group of pitchers is due to differences in the pitchers' probabilities and also due to chance binomial variation. Among all the various rates, we find that a strikeout rate is one of the most accurate estimates of the corresponding probability of a pitcher performing the associated task. We examine the distribution of strikeout, walk and runs-prevented "true" rates of pitchers across the years. By use of our model, we are able to judge the magnitude of a great strikeout season. A z-score statistic is used to rank the greatest strikeout seasons of baseball history and this ranking is contrasted with other traditional ways of ranking pitchers.

KEYWORDS: random effects model, strikeout rates, outliers

1. Introduction

Who was the greatest strikeout pitcher of all time? When we think of great strikeout pitchers, current pitchers such as Pedro Martinez, Randy Johnson, and Curt Schilling come to mind, and if we think back in history, Nolan Ryan and Bob Feller were both famous for their strikeout ability.

How can we measure a pitcher's propensity to strike out batters? Several obvious measures are currently used in ranking the best strikeout pitchers. The first measure is simply the total number of strikeouts. Table 1 lists the top 25 pitching seasons in baseball history by the number of strikeouts. From this perspective, Nolan Ryan had the best season in 1973 when he struck out 383 batters.

Table 1: Ranking of 25 best strikeout seasons using number of strikeouts.

	PLAYER		YEAR	SO
1.	Nolan Ryan		1973	383
2.	Sandy Koufax		1965	382
3.	Randy Johnson		2001	372
4.	Nolan Ryan		1974	367
5.	Randy Johnson		1999	364
6.	Rube Waddell		1904	349
7.	Bob Feller		1946	348
8.	Randy Johnson		2000	347
9.	Nolan Ryan		1977	341
10.	Randy Johnson		2002	334
11.	Nolan Ryan		1972	329
12.	Randy Johnson		1998	329
13.	Nolan Ryan		1976	327
14.	Sam McDowell		1965	325
15.	Curt Schilling		1997	319
16.	Sandy Koufax		1966	317
17.	Curt Schilling		2002	316
18.	J.R. Richard		1979	313
19.	Pedro Martinez		1999	313

20.	Walter	Johnson	1910	313
21.	Steve	Carlton	1972	310
22.	Mickey	Lolich	1971	308
23.	Randy	Johnson	1993	308
24.	Mike	Scott	1986	306
25.	Sandy	Koufax	1963	306

What's wrong with using the total number of strikeouts as our measure? First, this measure excludes certain type of pitchers, such as relievers, who have relatively few opportunities to strike out batters. The list of best pitchers with respect to the number of strikeouts will consist only of starters who pitch for many innings. Second, one pitcher may have more strikeouts than a second pitcher simply because he has faced more batters. We should adjust the number of strikeouts for the number of batters faced.

The simplest adjustment is to consider a pitcher's strikeout rate defined as the number of strikeouts (SO) divided by the number of batters faced (BFP):

$$\text{SO rate} = \text{SO} / \text{BFP}.$$

This measure is similar to the more common Strikeouts per Nine Innings (SO - 9). Our SO rate is simply is the fraction of batters faced who were struck out, while SO-9 tells us, on average, how many batters are struck out per inning. The measure SO-9 actually favors pitchers who tend to walk or allow hits to batters. Table 2 lists the top 25 pitching SO rate seasons for all time. (We only considered pitchers who faced at least 100 batters in a season to exclude pitchers who were not used regularly.) Note that this list looks dramatically different from the first list since it contains many relief pitchers, more specifically closers, who are very able to strike out hitters in the few innings that they pitch during a game.

Table 2. Ranking of 25 best strikeout seasons using strikeout rate (pitchers with at least 100 BFP were included).

PLAYER		YEAR	SO	BFP	SO RATE	
1.	Eric	Gagne	2003	137	306	44.8
2.	Billy	Wagner	1999	124	286	43.4
3.	Brad	Lidge	2004	157	369	42.5
4.	Armando	Benitez	1999	128	312	41.0
5.	Billy	Wagner	1998	97	247	39.3
6.	Rob	Dibble	1992	110	286	38.5
7.	Billy	Wagner	1997	106	277	38.3
8.	Pedro	Martinez	1999	313	835	37.5
9.	Randy	Johnson	2001	372	994	37.4
10.	Rob	Dibble	1991	124	334	37.1
11.	Francisco	Rodriguez	2004	123	335	36.7
12.	Eric	Gagne	2002	114	314	36.3
13.	Robb	Nen	2000	92	256	35.9
14.	Tom	Gordon	2001	67	187	35.8
15.	Dan	Plesac	2001	68	190	35.8
16.	Pedro	Martinez	2001	163	456	35.7
17.	Rob	Dibble	1990	136	384	35.4
18.	Tom	Henke	1987	128	363	35.3
19.	Rob	Dibble	1989	141	401	35.2
20.	Eric	Gagne	2004	114	326	35.0
21.	Armando	Benitez	2000	106	304	34.9
22.	Matt	Mantei	1999	99	284	34.9
23.	Jose	Valverde	2003	71	204	34.8
24.	Pedro	Martinez	2000	284	817	34.8
25.	Byung-Hyun	Kim	2000	111	320	34.7

This brief analysis raises several preliminary questions.

- Should we distinguish starters from relievers when we consider strikeout ability?
- Scanning Table 1 and Table 2, we note that modern pitchers dominate the lists.

Is this right?

First, most fans would likely want to distinguish starters from relievers with respect to strikeout rates. It seems more difficult for a starter to strike out batters than a reliever since the starter has to pitch many innings during a game. In contrast, some relievers may face only a few batters in a game and (due to the game situation) there may be a strong desire to strike out individual batters. Moreover, relievers are often brought into a game for a favorable matchup with a particular hitter. Second, it is unfair to say that modern pitchers have the most talent to strike out batters. Over the years, the game of baseball has gone through many changes that have had a major impact on baseball records. Specifically, changes in the pitching mound, strike zone, and players' propensity to strike out would have a significant effect on the number of strikeouts. For this reason, it would seem appropriate to make adjustments for strikeout rates for the players' eras.

The first goal of this paper is to take a historical look at strikeout rates over the years to get a basic understanding how the pattern of strikeouts have changed from the early 1900's. Once this basic pattern has been established, the greatness of a particular pitcher's strikeout season can be judged by comparing his pitching statistics to the distribution of statistics of all pitchers for that season.

We will consider a number of different count statistics used in evaluating pitchers. The goal of pitching is to prevent runs from being scored and the best pitchers are obviously the players who allow the fewest runs. But there are other aspects of a pitching performance such as striking out or walking hitters that are related to the primary goal of preventing runs scored. If we consider a particular statistic, such as a count of strikeouts, then we are interested in the pitcher's true strikeout rate defined to be his probability of striking out a batter. Using this definition, a player has many true rates, such as a strikeout probability, a walk probability, and a runs-allowed probability. We wish to estimate the collection of

pitcher probabilities based on data for a particular season and then see how the distribution of probabilities has changed over the history of baseball.

A number of authors have commented on the change in pitching effectiveness over the years of professional baseball and the reasons for the change. For example, James (2001) describes the changes in the strike zone and the height of the pitchers mound in some ballparks in the 1960s that contributed to the domination by the pitchers in this era. Since the characteristics of pitching and hitting have changed over the years, Thorn and Palmer (1985) were one of the earlier authors to argue for a “theory of relativity” where a statistic such as a batting average is viewed in the context of the year, say by normalizing the batting average by the league average. The role of the standard deviation in this adjustment was described by Gould (1983) who argued why it is very unlikely for a hitter in the future to have a .400 batting average. Schell (1999, 2005) is one of the most sophisticated efforts to adjust hitting statistics with respect to ballparks, talent pools, and late-career declines.

The ability of a pitcher to strike out a batter may change as a function of his age. Different authors have explored the pattern of career trajectories of player performances. James (1982) presented his “Theory of Aging” and suggested the ages where ballplayers reach their peak performance. Berry et al (1999) and Albert (2002) describe different statistical approaches for making inferences about career trajectories of players.

A second important theme given less attention in the literature is the role of luck or chance variation in baseball statistics. James et al (1993) describe the role of chance variation in baseball competition and distinguished a team’s talent from its performance in a season. One of the major themes in Albert and Bennett (2001) is learning about the talents of players and teams on the basis of their performances. One general conclusion from Albert and Bennett (2001) is that

chance variation plays a large role in team competition and season batting statistics.

Sections 2 and 3 consider a number of statistics for pitchers including wins, losses, runs allowed, earned runs allowed, hits, walks and strikeouts. Among all of these different count statistics, it is shown that the observed rate of strikeouts is the single rate statistic that most accurately estimates the pitcher's underlying probability or true rate of that event. In other words, if one ranks all of the different pitching statistics with respect to the percentage of the total variation of the pitching data due to difference in probabilities (rather than chance variation), then the strikeout rate is the top statistic.

Section 4 takes a historical look at three basic pitching statistics – strikeouts, walks, and runs. For each season, we find the distribution of probabilities of the pitchers for striking out batters, for walking batters, and for giving up runs. We will see significant changes in these “random-effects” distributions across years, both in the average and in the spread of the pitcher probabilities. We use this work to help find the greatest strikeout pitchers of all time. In Section 5, we use a z statistic to measure the magnitude of a particular strikeout season with respect to the distribution of strikeouts for all pitchers that season. We create new lists of the greatest strikeout seasons for starters and relievers that drastically modify the naïve lists and give recognition to some of the great strikeout seasons in the past. In Section 6, we look more carefully at the career trajectories of strikeout rates for several great pitchers in baseball history.

2. Review of measures to rate pitching.

All of the statistics used to rate pitchers are based on counts. We record the number of wins (W), the number of losses (L), the innings pitched (IP), the runs allowed (R), the earned runs allowed (ER), the number of walks (BB), the number of strikeouts (SO), the number of batters hit by a pitch (HBP), the number

of wild pitches (WP), the number of hits allowed (H), and the number of home runs allowed (HR). All of these counts are observed in a context, that is, the number of innings pitched. To avoid fractions of innings and avoid the complications due to plays that are not outs, we use the number of batters that a pitcher faces (BFP). In the following, we won't consider the number of wins and losses of a pitcher since it is well known that these counts are confounded with the ability of the pitcher's team. Also, we do not consider wild pitches and batters hit by a pitcher, since these are relatively rare events for most pitchers.

Suppose we observe the number of runs allowed and batters faced for all pitchers in a particular season and we compute the run rates R/BFP . We will notice a lot of variation in these data. What is the cause of this variation? As described by Albert (2005), there are two principal causes for the variation in the pitcher run rates:

- Variation in the underlying pitcher probabilities. Basically, some pitchers are more talented in preventing runs than other pitchers, where talent is measured by the pitcher's probability of allowing a batter to score.
- Chance variation. There is a chance aspect of giving up runs that is beyond a pitcher's ability. Fans have seen innings where a pitcher is victimized by a number of bloop hits or bad plays by fielders and the opposing team scores runs. The variation due to these other causes outside a pitcher's control is called chance variation and will be modeled by a binomial distribution.

The main point of this paper is to separate out the two types of variation. Given the data on runs and batters faced by all pitchers in a particular season, we would like to quantify exactly how much of the variation in the observed run rates is due to the variation in pitcher probabilities and how much is due to binomial chance variation.

The two sources of variation can be estimated by fitting the following random effects model. The model is described in three stages.

Stage 1. Let y_i and n_i denote respectively the observed count and the number of opportunities for the i th pitcher. In our runs example, y_i would denote the number of runs allowed and n_i the number of batters faced for the i th pitcher. We assume that y_i has a binomial (n_i, p_i) distribution where p_i is the pitcher's probability – in our example, p_i would be the probability that the i th pitcher allows a batter to score a run. The definition of this probability implicitly ignores the variation of the batters' skill in not striking out. One can interpret p_i as the probability that the pitcher strikes out a batter of average skill.

Stage 2. The pitcher probabilities p_1, \dots, p_N , where N is the number of pitchers, is assumed to be a random sample from an *talent distribution* $g()$ – since the probabilities fall between 0 and 1, it is convenient to let the talent distribution be a beta distribution with parameters a and b .

Stage 3. Since the parameters a and b of the beta distribution are unknown, they are assigned a noninformative prior distribution. To describe this distribution, it is helpful to reparameterize a and b to the mean parameter $m = \frac{a}{a+b}$ and the precision parameter $K = a+b$. We assume m and K are independent with m assigned the prior $(m(1-m))^{-1}$ and K the prior proportional to $(1+K)^{-2}$

We fit this random effects model to data for all pitchers for a particular season. The beta parameters a and b are estimated by the mode (\hat{a}, \hat{b}) of the

posterior distribution of (a, b) . The pitcher probability p_i can be estimated by the approximate posterior mean

$$\hat{p}_i = \frac{y_i + \hat{a}}{n_i + \hat{a} + \hat{b}}.$$

By inspecting the fitted values \hat{a} and \hat{b} we learn about the range of the talent distribution of pitchers. In addition, it allows us to determine what fraction of the pitching data is due to differences in probabilities and what fraction is due to binomial variation. By comparing the fitted models for different measures, we can see which pitcher measures are primarily talent-driven and what measures are chance-driven. The estimate of the pitcher probability \hat{p}_i moves the observed rate y_i/n_i towards the mean estimate $\hat{a}/(\hat{a} + \hat{b})$.

Albert (2005) used this model to look at batting data for a particular season. What he learned was that certain statistics such as a batting average were mainly chance-driven, and other hitting statistics such as a strikeout rate and walk rate were more related to the talent of a batter. The outcome that a ball put in-play falls in for a hit is primarily driven by binomial variation. It would be reasonable to think that these conclusions for batters would also apply to similar data for pitchers.

3. Fitting of the random effects model to pitching statistics data

The random effects model described in Section 2 was fit to data for all pitchers in the 2003 season. We considered the following seven pitching statistics:

- *Walk rate*: Number of walks (BB) divided by batters faced pitcher (BFP).
- *Strikeout rate*: Number of strikeouts (SO) divided by batters faced pitcher (BFP).

- *Home run rate:* Number of home runs allowed divided by balls put in play (IP), where $IP = BFP - BB - HBP$.
- *Non-home run rate:* Number of non-home run hits allowed divided by balls put in play.
- *Run rate:* Number of runs (R) scored divided by batters faced pitcher (BFP).
- *Earned run rate:* Number of earned runs (R) scored divided by batters faced pitcher (BFP).

Table 3 presents the estimated values \hat{a} and \hat{b} of the beta parameters for all seven statistics. This table also shows the mean and standard deviation of each talent distribution and Figure 1 displays the middle 90% intervals for the probabilities of each type of event.

Table 3: Estimated parameters of the talent distribution for each of the seven pitching statistics based on data from the 2003 season.

STATISTIC	a	b	a+b	mean	stand. dev.
walk	12.1	120.8	132.9	0.091	0.025
strikeout	10.6	53.9	64.5	0.164	0.046
in-play home runs	24.0	608.0	632.1	0.038	0.008
in-play hits					
not hr	276.6	710.7	987.3	0.280	0.014
runs	15.7	107.1	122.8	0.128	0.030
earned runs	14.3	106.4	120.6	0.118	0.029
batting average	77.6	219.6	297.3	0.261	0.025

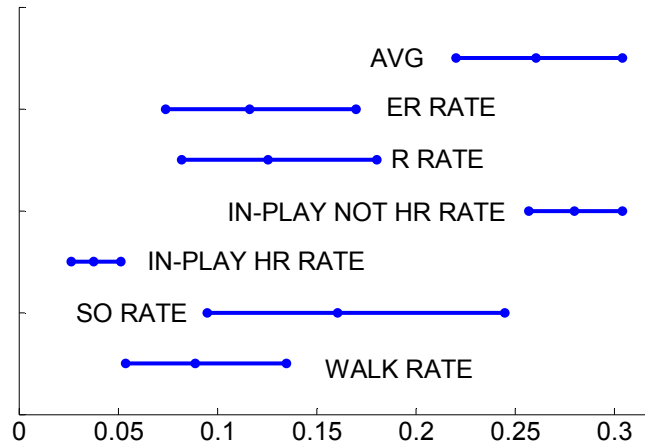


Figure 1: Graphs of middle 90% of talent distributions for seven pitching statistics based on data from the 2003 season.

The sum of the estimated beta parameters, $\hat{a} + \hat{b}$, is a measure of the precision of the talent distribution of the pitching rates. Small values of $\hat{a} + \hat{b}$ correspond to a wide talent distribution – in this case, there are substantial differences in the pitcher probabilities. In contrast, large values of $\hat{a} + \hat{b}$ correspond to statistics where there are subtle differences in the pitchers’ probabilities. By using the estimated values of $\hat{a} + \hat{b}$ given in Table 3, we can sort the seven statistics on the talent/chance dimension shown below. Since the estimated values of $\hat{a} + \hat{b}$ for runs and earned runs are so similar and not significantly different, we place them together in the ordering.

STRIKEOUTS > EARNED RUNS, RUNS > WALKS > AVG > IN-PLAY HR > IN-PLAY HIT NOT HR

MORE TALENT ----- MORE CHANCE

These results mirror the results on batting data reported in Albert (2005). Pitchers have varying probabilities to strike out batters, allow runs and walk batters. In contrast, pitchers' in-play batting average and batting average (against) are much more influenced by chance variation. This confirms that it is worthwhile to look at strikeout data, since the strikeout rate is one of the most "talent-driven" pitching statistics. It might be surprising that the in-play home run rate is categorized as a chance statistic. One can think of several pitchers such as Curt Schilling and Eric Milton that seem to let up a lot of home runs. But it should be emphasized that we are looking at the home run rates for all pitchers and this says that most of the variability in these home run rates is due to chance – it doesn't exclude the possibility of several outliers like Schilling and Milton.

One way of describing the relative sizes of the talent and chance sources of variations is by a "percentage of chance" statistic. The variance of a (beta) talent distribution with parameters a and b is given by

$$\text{Var}(\text{talent}) = \frac{m(1-m)}{a+b+1},$$

where m is the mean $m = a/(a+b)$. If we have a pitcher with a probability m who faces n batters, then the variance of his observed rate due to chance is given by the binomial formula

$$\text{Var}(\text{luck}) = \frac{m(1-m)}{n}.$$

The percentage of the total variation that is due to chance is defined by

$$\text{Percentage}(\text{chance}) = 100 \times \frac{\text{Var}(\text{chance})}{\text{Var}(\text{talent}) + \text{Var}(\text{chance})}.$$

Table 4 presents the percentages of chance that is involved with different pitching statistics based on the estimated values of the beta parameters a and b . Since the chance percentage is dependent on the observed number of batters that a pitcher faces, the percentage is given for BFP = 200, 500, and 1000. To understand these numbers, suppose one has a collection of in-play not-home runs rates for a group of pitchers, each with BFP = 200. By Table 4, 83% of the variation in these rates is due to chance variation and only 17% is due to different probabilities between the pitchers. In contrast, suppose we collect the strikeout rates for a group of pitchers with BFP = 200. Since a strikeout rate is driven more by a pitcher's probability, only 25% of the variation in these strikeout rates is due to chance and a large 75% of the variation is due to differences in the pitchers' probabilities to strike out batters. If we collect strikeout rates for a group of pitchers each with 1000 BFP (instead of 250 BFP), then practically all (94%) of the variation is due to pitcher probabilities and only 6% is due to chance.

Table 4: Percentage of chance in various pitcher rate statistics for different number of batters faced pitcher (BFP).

STATISTIC	BFP=200	BFP=500	BFP=1000
walk	40	21	12
strikeout	25	12	6
in-play home runs	76	56	39
in-play hits not hr	83	66	50
runs	38	20	11
earned runs	38	20	11
batting average	60	37	23

4. Talent distributions through the years

To get a historical perspective on strikeout, walk and run-preventing probabilities, we fit the random effects model for all pitchers for all seasons from 1871 through 2005. For each season and statistic, we summarized the fit by the beta parameters \hat{a} and \hat{b} . Figures 2, 4, and 6 plot the fitted talent distributions for strikeouts, walks, and runs respectively using line graphs for each year. The middle of the line graph is the median of the talent distribution and the endpoints are the 25th and 75th percentiles. These figures are helpful in understanding the locations of the distributions and how the typical probability of a pitcher changes over time. Figures 3, 5 and 7 plot the standard deviations of the talent distributions across time. These plots are helpful in seeing how the spread or variation in pitcher probabilities has changed over years. Smooth curves (found using the loess algorithm) are placed on these standard deviation graphs to show the basic patterns.

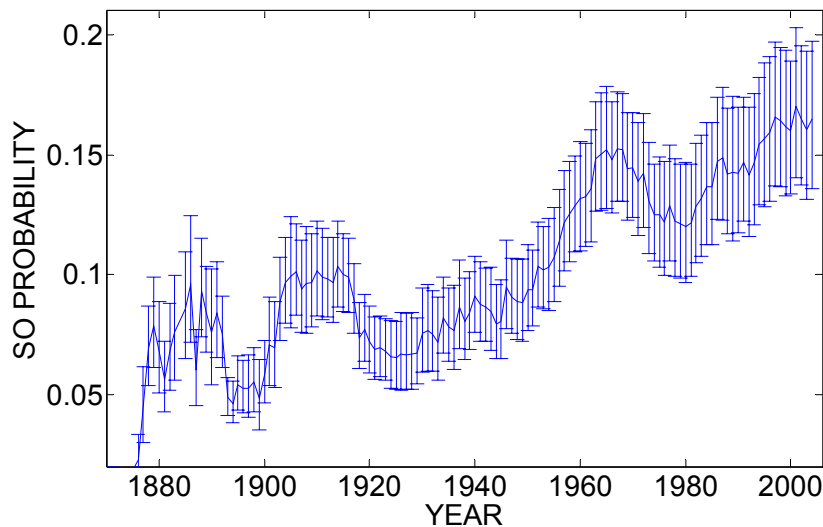


Figure 2: Fitted talent distributions for strikeouts over baseball history. Each bar displays the lower quartile, median, and upper quartile of the distribution of strikeout probabilities.

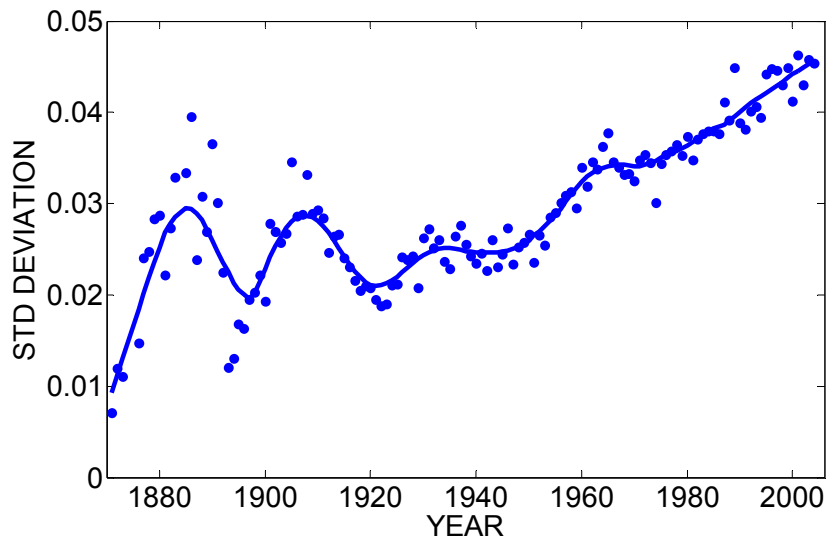


Figure 3: Standard deviations of talent distributions for strikeouts over baseball history. The smooth curve is a loess fit to show the basic pattern of the scatterplot.

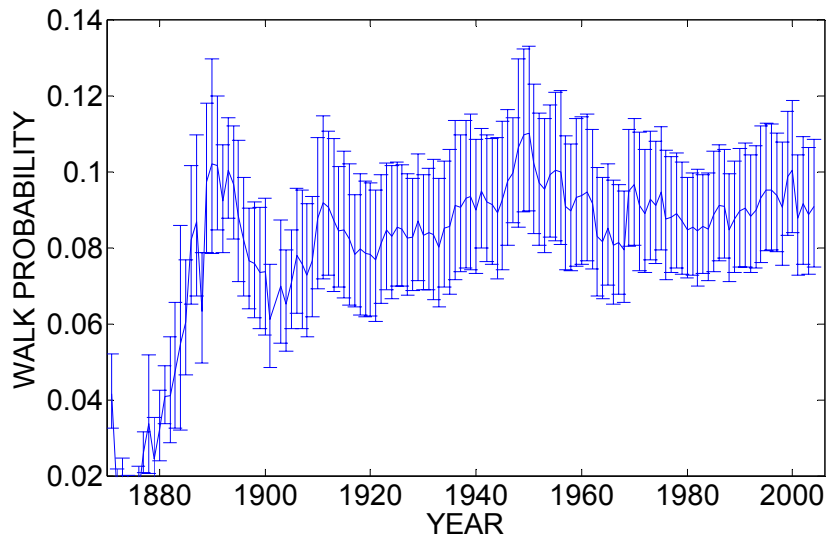


Figure 4: Fitted talent distributions for walks over baseball history. Each bar displays the lower quartile, median, and upper quartile of the distribution of walk probabilities.

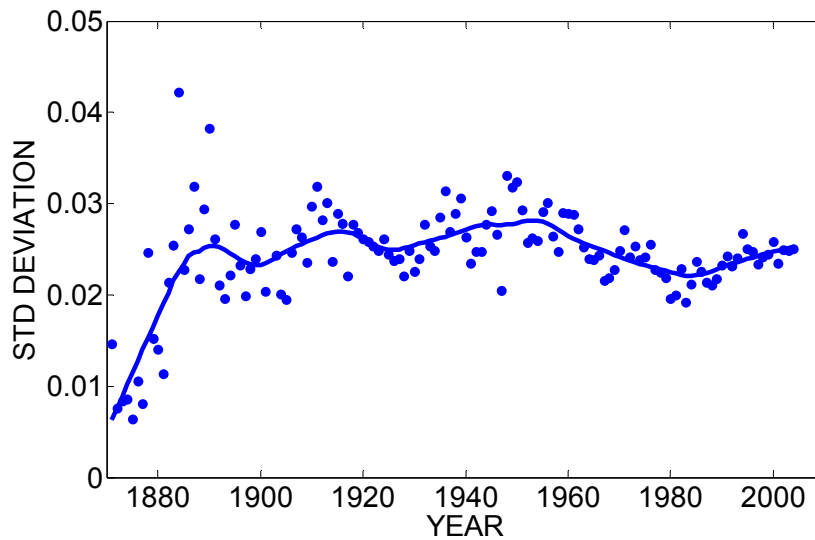


Figure 5: Standard deviations of talent distributions for walks over baseball history. The smooth curve is a loess fit to show the basic pattern of the scatterplot.

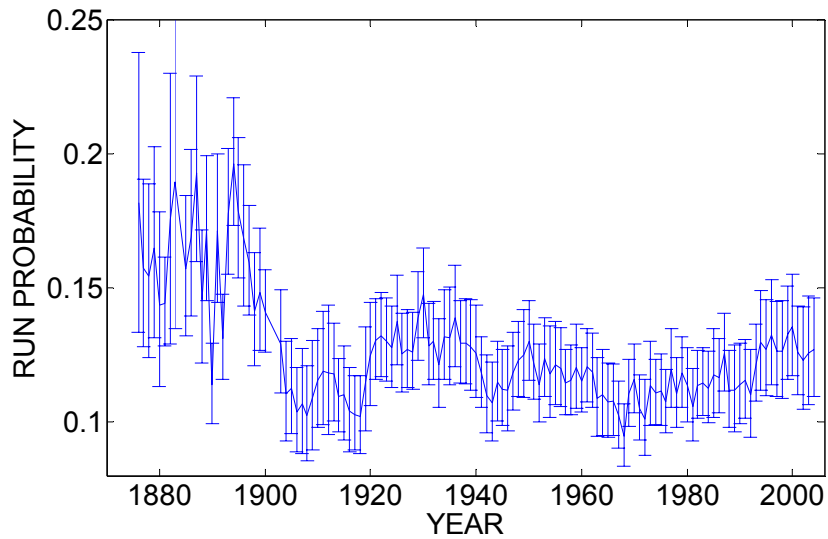


Figure 6: Fitted talent distributions for runs over baseball history. Each bar displays the lower quartile, median, and upper quartile of the distribution of run probabilities.

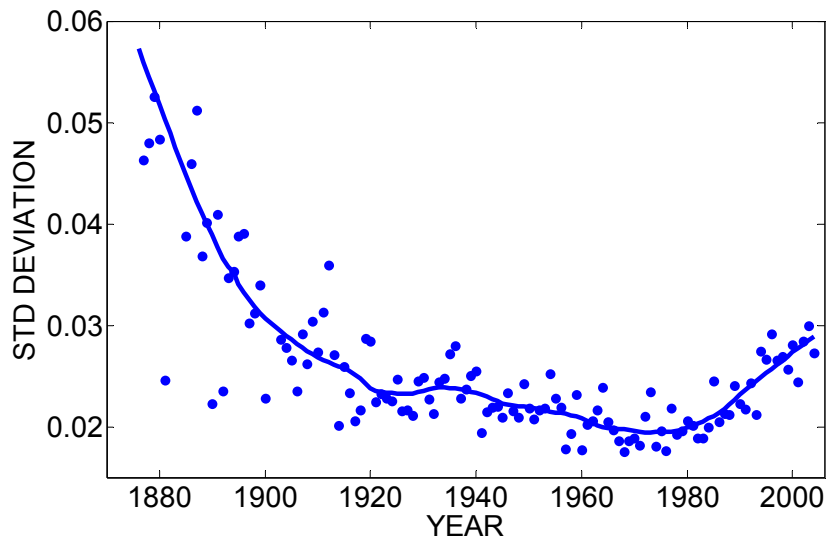


Figure 7: Standard deviations of talent distributions for runs over baseball history. The smooth curve is a loess fit to show the basic pattern of the scatterplot.

There are some interesting observations from inspecting these figures.

1. There have been dramatic differences in pitchers' probabilities to strike out batters in history. As Figure 2 indicates, pitchers' strikeout probabilities were variable in the early years, low in 1920, gradually increased until a peak in the mid 1960s, dropped to a low value in 1980, and have been steadily increasing in the last 25 years. Figure 3 shows that the spread in pitchers' probabilities to strike out batters has steadily increased since 1920. So currently, pitchers are striking out a lot of batters and there is much variation in the pitcher probabilities.
2. In contrast, the pitcher's probabilities in walking hitters have stayed relatively constant across time. The standard deviation plot for walks

(Figure 5) is interesting – since 1980, there has been increasing variation in the pitchers' walk probabilities.

3. Figure 6 shows much up-and-down behavior in the pitcher probabilities to allow runs. Presumably, this pattern would mirror the changes in run production across years. Again the standard deviation graph in Figure 7 is interesting. The variation in pitcher probabilities to allow runs steadily decreased until about 1980 and has been steadily increasing in the last 25 years.

Figures 1 and 7 focused on the estimated talent distribution of probabilities given by the beta curve. To learn more about the pitcher strikeout probabilities, it is instructive to look at the set of estimated probabilities

$\left\{ \hat{p}_i = \frac{y_i + \hat{a}}{n_i + \hat{a} + \hat{b}} \right\}$ for several seasons. Figure 8 displays histograms of the

estimated strikeout probabilities for all pitchers for the seasons 1904, 1924, 1944, 1964, 1984, and 2004. We see graphically the changes in the location and spread of these distributions. For the years 1964, 1984 and 2004 the distributions are notably right-skewed. The estimated strikeout probabilities for several pitchers (Rube Waddell in 1904, Dazzy Vance in 1924, Dick Radatz in 1964, Dwight Gooden in 1984, Francisco Rodriguez and Brad Lidge in 2004) stand apart from the main group of estimated probabilities. When checking the adequacy of the model fit, it was found that there was some lack of fit with regards to these extreme outliers. For example, when checking the goodness of fit of the model to the 2004 data, it was found that Brad Lidge's strikeout rate was higher than one would predict from this random effects model. Last, note that the distribution of strikeout probabilities for the 2004 season shows an unusually heavy right tail – this could be due to the prevalence of the specialized reliever such as closers and set-up men in recent season. This last observation may help to explain the

pattern of increasing standard deviations of the strikeout talent distribution in Figure 3. In recent years there has an increase in the number of specialized relievers such as closers and set-up men who generally have higher strikeout rates than the starting pitchers. The disparity in the strikeout probabilities between these two groups of pitchers can cause the greater standard deviation.

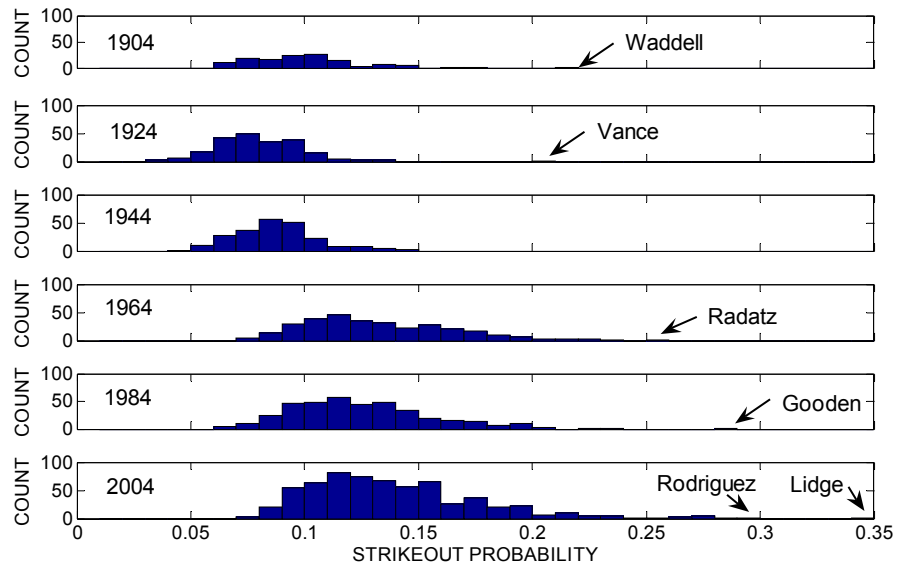


Figure 8. Histograms of estimated strikeout probabilities for all pitchers in six seasons. Pitchers with unusually large estimated probabilities are labeled.

5. Outstanding strikeout seasons

Figure 2 shows that there have been large differences in strikeout probabilities of pitchers in baseball history. If one was told that a pitcher strikes out 1/3 of the batters that he faces, this comment is really not useful on face value. One would have to know the context of this statistic. This statement would be much more impressive in 1920 when strikeouts were rare compared to the current day when strikeouts are more common.

We can use the fitted talent distributions to establish a context for a player's strikeout accomplishment. In 1976, Nolan Ryan struck out 327 of the 1196 batters he faced for a strikeout percentage of 27.3. Is this a high percentage? To assess this, we compute a z-score using the fitted strikeout talent distribution for the 1976 season in which Ryan played. This z-score will tell us the relative standing of Ryan's strikeout percentage.

For our fitted random effects model, it can be shown that an observed rate y_i/n_i has an estimated mean $\hat{m} = \hat{a}/(\hat{a} + \hat{b})$ and estimated standard

deviation $s = \sqrt{\frac{\hat{m}(1-\hat{m})}{n_i} + \frac{\hat{m}(1-\hat{m})}{\hat{a} + \hat{b} + 1}}$. So the z-score

$$z = \frac{y_i/n_i - \hat{m}}{s}$$

is a useful way of gauging the size of an observed rate. For the 1976 season, pitchers who faced 1196 batters would have strikeout rates with a mean of $m = .125$ and a standard deviation of $s = .0367$. The z-score for Ryan's strikeout rate of .273 would be

$$z = \frac{.273 - .125}{.0367} = 4.05.$$

This z-score means that Ryan's strikeout count is approximately four standard deviations above the average strikeout rate for a pitcher facing the same number of batters that particular season.

This z statistic is a straightforward way of adjusting a pitcher's strikeout rate for the season in which he played. We can rank pitcher accomplishments by the z statistics – a pitcher with the highest z score is the player who did best relative to his peers for that particular season.

6. The adjusted best strikeout pitchers since 1900

We present several tables of strikeout leaders using our method of adjusting strikeout totals for the pitchers' seasons. Table 5 presents the best seasons for all starting pitchers who faced at least 100 batters in a season. Table 6 presents the best strikeout seasons for pitchers who were primarily relievers. Note that the top three seasons among starters are held by two pitchers, Dazzy Vance and Bob Feller, who pitched during the 1920's, 1930's, and 1940's when strike out rates were relatively low. The pitchers with more than one season on this list are Dazzy Vance (4 seasons), Bob Feller (3 seasons), and Rube Waddell (2 seasons). Note that Dazzy Vance in 1923 struck out only 16.6% of the batters he faced, but this accomplishment was very remarkable (z score of 4.72) in a season where strikeouts were not very common. Among relievers, the top strikeout seasons are achieved by modern relief pitchers such as Eric Gagne, Billy Wagner and Brad Lidge who specialized as closers.

Table 5. Ranking of best 20 strikeout seasons of starting pitchers by z-score that judges the performance of the strikeout season relative to the strikeout abilities of pitching for that season.

rank	name	year	so	bfp	rate	z
1	Dazzy Vance	1924	262	1221	21.5	6.58
2	Bob Feller	1936	76	279	27.2	6.22
3	Dazzy Vance	1925	221	1089	20.3	6.02
4	Rube Waddell	1903	302	1307	23.1	5.20
5	Dazzy Vance	1926	140	713	19.6	4.88
6	Hal Newhouser	1946	275	1176	23.4	4.77
7	Bob Feller	1937	150	651	23.0	4.74
8	Dazzy Vance	1923	197	1187	16.6	4.72
9	Johnny Vander Meer	1941	202	945	21.4	4.72

10	Bob Feller	1946	348	1512	23.0	4.69
11	Bill Bailey	1914	131	536	24.4	4.67
12	Herb Score	1955	245	978	25.1	4.59
13	Rube Waddell	1904	349	1548	22.5	4.56
14	Pedro Martinez	1999	313	835	37.5	4.49
15	Herb Score	1956	263	1022	25.7	4.44
16	Dwight Gooden	1984	276	879	31.4	4.40
17	Lefty Grove	1926	194	1072	18.1	4.39
18	Bob Feller	1939	246	1243	19.8	4.37
19	Sandy Koufax	1957	122	444	27.5	4.36
20	Nolan Ryan	1974	367	1392	26.4	4.35

Table 6. Ranking of best 20 strikeout seasons of relievers by z-score that judges the performance of the strikeout season relative to the strikeout abilities of pitching for that season.

rank	name	year	so	bfp	rate	z
1	Eric Gagne	2003	137	306	44.8	5.63
2	Billy Wagner	1999	124	286	43.4	5.37
3	Rob Dibble	1992	110	286	38.5	5.31
4	Brad Lidge	2004	157	369	42.5	5.20
5	Rob Dibble	1991	124	334	37.1	5.17
6	Armando Benitez	1999	128	312	41.0	4.94
7	Rob Dibble	1990	136	384	35.4	4.88
8	Ryne Duren	1959	96	322	29.8	4.79
9	Bruce Sutter	1977	129	411	31.4	4.62
10	Billy Wagner	1998	97	247	39.3	4.59
11	Skip Lockwood	1975	61	192	31.8	4.53
12	Tom Henke	1987	128	363	35.3	4.44
13	Bill Wagner	1997	106	277	38.3	4.28
14	Rob Dibble	1989	141	401	35.2	4.26
15	Ryne Duren	1958	87	307	28.3	4.25

16	Tom Hall	1970	184	619	29.7	4.24
17	John Hiller	1975	87	295	29.5	4.24
18	Duane Ward	1993	97	282	34.4	4.23
19	Ron Davis	1981	83	285	29.1	4.19
20	Mark Littell	1978	130	447	29.1	4.17

To see what seasons are represented in these two lists of top strikeout pitchers, Figure 9 displays a dotplot of the seasons. Note that the best starting strikeout pitchers are heavily represented from the years 1920 through 1960. In contrast, the best strikeout relievers are mainly from recent years – this is expected since specialized relief pitchers such as closers and set-up men are modern phenomena.

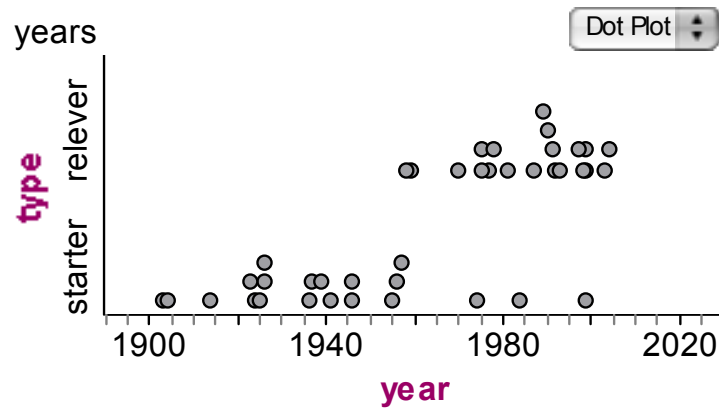


Figure 9. Dotplot of season of the top 20 starters and relievers with the best adjusted strikeout rates since 1900.

7. Taking a closer look at some great strikeout pitchers

To further illustrate the talent dimension of striking out batters, Figure 10 displays a scatterplot of the strikeout rates of all pitchers who faced at least 500 batters for both the 2003 and 2004 seasons. Note that there is a strong association

between the 2003 and 2004 rates and a pitcher's strikeout rate for one season is a good predictor of his rate for the following season. Horizontal and vertical lines in the plot are drawn at a value of .25 for each year, and the labeled points correspond to pitchers who struck out at least 25% of batters faced for one of the two seasons. The 14 pitchers identified in this figure include some famous strikeout artists who are finishing their careers (Johnson, Martinez, Schilling and Clemens) together with some young pitchers (Prior, Wood, Santana, Beckett, Sheets).

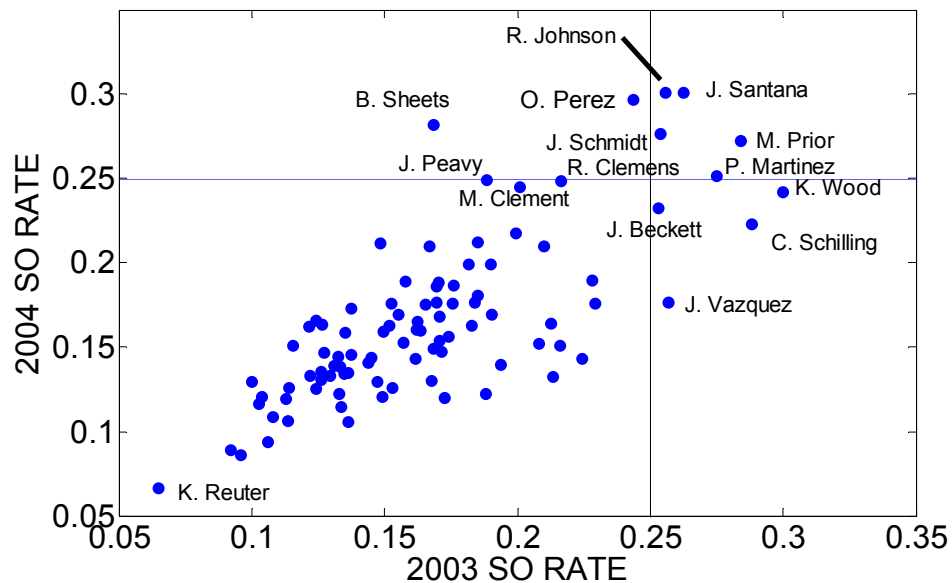


Figure 10. Scatterplot of strikeout rates of all pitchers for the 2003 and 2004 seasons with at least 500 BFP.

Since a strikeout rate is a good estimate of the associated talent, it is interesting to observe the career trajectory of a pitcher's strikeout performance as a function of age. For a pitcher, one might expect his career trajectory to have a parabolic shape where the pitcher matures and peaks in strikeout performance during mid-career and then declines until his retirement. We observe this

parabolic shape for some pitchers, but other pitchers display a constant strikeout pattern over their careers. We illustrate these phenomena by displaying the career trajectory of some Hall-of-Fame pitchers.

The career trajectories of Nolan Ryan, Bob Feller, Sandy Koufax, Bob Gibson, and Steve Carlton are displayed in Figures 11-15. For each season, we have adjusted a pitcher's strikeout rate by use of the z-score method described in the previous section. So these graphs have removed the effects due to different talent distributions across years. Note that Ryan and Koufax displayed remarkably high strikeout rates during the last years of their careers. In contrast, Feller peaked earlier and declined steadily after he returned home from World War II. Gibson and Carlton peaked in their middle 30's and declined in strikeout probability during the last years of their careers.

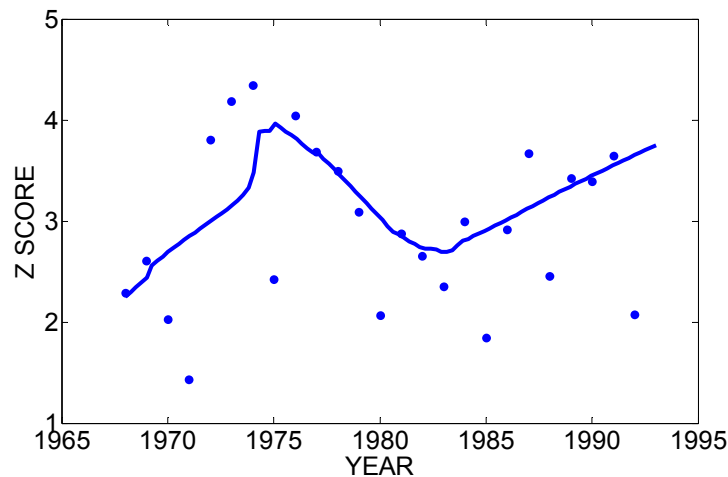


Figure 11: Career trajectory of the adjusted strikeout rates (z-scores) of Nolan Ryan over his career. A loess smoothing curve is placed on top to show the basic pattern.

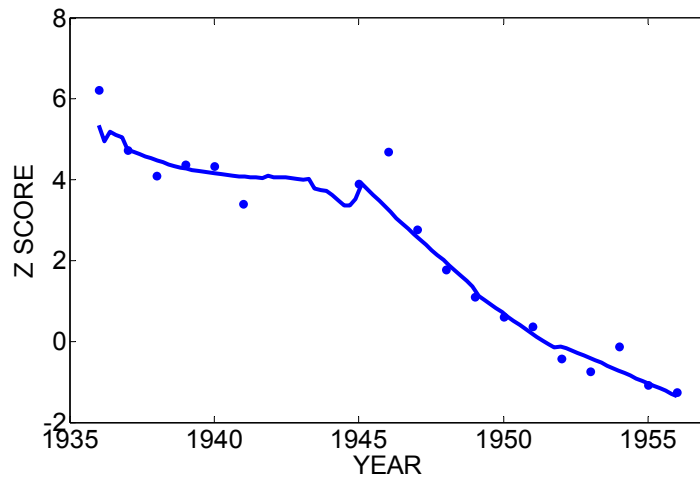


Figure 12: Career trajectory of the adjusted strikeout rates (z-scores) of Bob Feller over his career. A loess smoothing curve is placed on top to show the basic pattern.

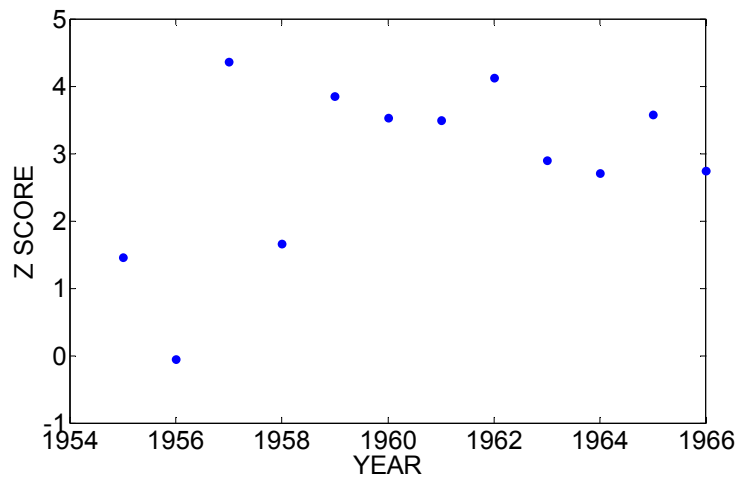


Figure 13: Career trajectory of the adjusted strikeout rates (z-scores) of Sandy Koufax over his career. A loess curve was not found helpful in identifying the pattern in the graph.

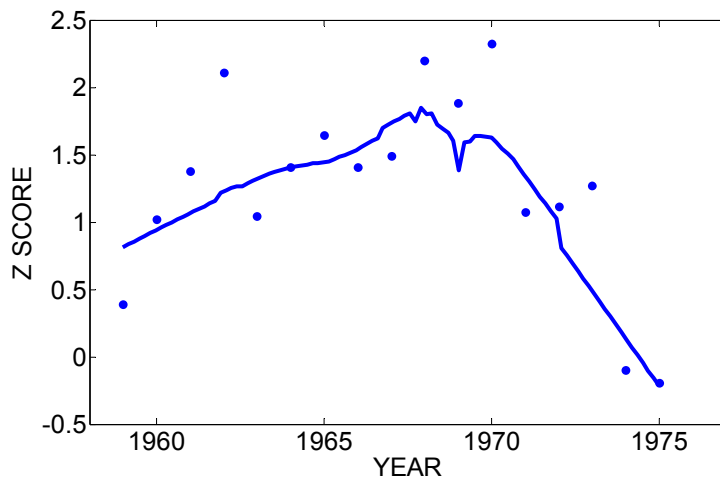


Figure 14: Career trajectory of the adjusted strikeout rates (z-scores) of Bob Gibson over his career. A loess smoothing curve is placed on top to show the basic pattern.

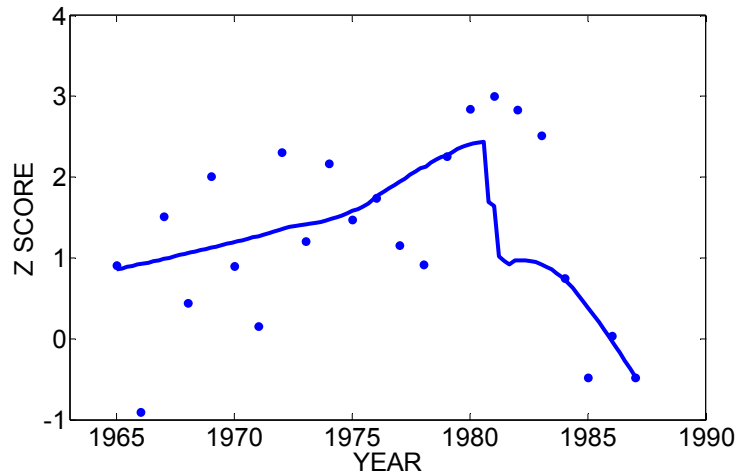


Figure 15: Career trajectory of the adjusted strikeout rates (z-scores) of Steve Carlton over his career. A loess smoothing curve is placed on top to show the basic pattern.

8. Final remarks

There have been some efforts by sabermetricians to properly adjust a pitcher's strikeout counts for the era in which he played. One idea described by Treder (2004) is to compute the ratio of a pitcher's strikeout rate to the league average, obtaining a "league-normalized" strikeout rate denoted by K^+ . The top season performances with regard to the K^+ statistic are similar to the top performances reported in Table 5.

There are several attractive features of our approach for evaluating pitchers. First, by use of the random effects model, one is able to separate out the variation in the pitching data due to the difference in probabilities and the variation due to binomial chance variation. By examining the average and spreads of the fitted talent distributions across years, it is easy to see the change in pitching probabilities. One interesting aspect of this study is that there has been a notable increase in the variation of strikeout, walk, and runs-allowed pitcher probabilities in recent years, suggesting that there is currently a wide range of talent among major league pitchers. Second, we have suggested a general method for assessing the extremeness of great pitching accomplishments like a large number of strikeouts. The z -measure that we use adjusts for the average and spread of the pitching talent distribution. In contrast, the K^+ statistic only adjusts for the average pitching measure for a particular season. As Section 7 indicates, it is very interesting to explore a pitcher's career trajectory of strikeout rates, and pitchers appear to display different career trajectories in their strike out rates over their careers. For future work, it would be interesting to classify pitchers by the patterns of their career trajectories.

A related issue to strikeouts is the degree of control a pitcher has in preventing base hits. McCracken (2001) wrote an influential paper in which he said that generally pitchers had no control on the result of a ball that was placed in-play. His conclusions are substantiated in Section 3 where we found little

differences in the true pitcher probabilities of a batted ball falling in for a hit. Tippett (2003) in a large investigation shows that a few pitchers do have the talent to prevent base hits on balls in-play. By looking at the performances of pitchers with respect to striking out batters, walking batters, preventing home runs and preventing hits from balls in-play, and comparing these values with average performances, one can understand better why particular pitchers are successful.

References:

- Albert, J. (2002), "Smoothing Career Trajectories of Hitters", *By the Numbers* (Newsletter of the Statistical Analysis Committee of the Society of American Baseball Research), Vol 12, 3, 9-19.
- Albert, J. (2005), "A Baseball Hitter's Batting Average: Does It Represent Ability or Luck?", *Stats*, 42.
- Albert, J. and Bennett, J. (2003), *Curve Ball*, Springer.
- Albert, J., James, B., and Stern, H. (1993), "Answering questions about baseball using statistics", *Chance*, 6, 17-22.
- Berry, Scott M., Reese, C. Shane, and Larkey, Patrick D. (1999), "Bridging Different Eras in Sports", *JASA*, 94, 447, 661-686
- Gould, S. J. (1983), "Losing the Edge: The Extinction of the .400 Hitter." *Vanity Fair*, 120, 264-78.
- James, B. (1982), *The Bill James Baseball Abstract*, Balletine books.
- James, B. (2001), *The New Bill James Historical Baseball Abstract*, The Free Press.
- McCracken, V. (2001), "Pitching and Defense: How Much Control Do Hurlers Have?", *Baseball Prospectus* (www.baseballprospectus.com).
- Schell, M. (1999), *Baseball's All-Time Best Hitters*, Princeton University Press.

- Schell, M. (2005), *Baseball's All-Time Best Sluggers : Adjusted Batting Performance from Strikeouts to Home Runs*, Princeton University Press.
- Tippett, T. (2003), "Can Pitchers Prevent Hits on Balls in Play?", article on www.diamond-mind.com.
- Thorn, J. and Palmer, P. (1985), *The Hidden Game of Baseball*, Doubleday.
- Treder, S. (2004), "Strike Zone Dominance in Context (Dazzy and Pedro!)," *The Hardball Times* (www.hardballtimes.com).