



Using Statistical Process Control Charts to Identify the Steroids Era in Major League Baseball: An Educational Exercise

[Stephen E. Hill](#)

[Shane J. Schvaneveldt](#)

Weber State University

Journal of Statistics Education Volume 19, Number 1 (2011),
www.amstat.org/publications/jse/v19n1/hill.pdf

Copyright © 2011 by Stephen E. Hill and Shane J. Schvaneveldt all rights reserved. This text may be freely shared among individuals, but it may not be republished in any medium without express written consent from the authors and advance notification of the editor.

Key Words: Statistical Quality Control; Control Charts; Sports; Innovative Education.

Abstract

This article presents an educational exercise in which statistical process control charts are constructed and used to identify the Steroids Era in American professional baseball. During this period (roughly 1993 until the present), numerous baseball players were alleged or proven to have used banned, performance-enhancing drugs. Also observed during this period was an increase in offensive performance. In this exercise, students are given the opportunity to construct trial control limits from historical data, consider the presence of random and assignable cause variation, and analyze offensive performance for the 1993 to 2008 baseball seasons. From this analysis, students can consider the potential impact of performance-enhancing drugs on offensive performance in Major League Baseball.

1. Introduction

In this paper, an educational exercise is proposed that seeks to answer the question: “Could statistical process control techniques have been utilized to recognize the onset of the Steroids Era in Major League Baseball that is thought to have begun in 1993?” Students are given historical Major League Baseball batting average data and are asked to analyze the data, construct trial control chart limits, plot data from 1993 onwards against the control limits, and analyze the results.

This exercise is suitable for use in undergraduate or graduate-level courses in which a portion of the course is devoted to the teaching of statistical process control techniques including control

charts for individuals. The exercise is structured so that it can be presented as a small project, lab assignment, homework assignment, or as an in-class exercise or demonstration.

2. Background: Offensive Performance in Major League Baseball (1901-2008)

The “modern” history of Major League Baseball (the highest professional baseball league in the United States) is generally considered to have begun in 1901. Since this time, baseball has experienced a series of historical eras during which offensive performance was observed to be different than in other periods of baseball history. Major League Baseball’s eras of offensive performance as defined by [Schell \(2005\)](#), a renowned baseball statistician and writer, are used in this article. These eras are as follows: Deadball Era (1901-1919), Lively Ball Era (1920-1946), Post-World War II Era (1947-1962), Big Strike Zone Era (1963-1968), Designated Hitter Era (1969-1992), and Power Era (1993-Present). Each of these eras was unique, due to changes in rules or to other factors that altered the offensive production in the game. For example, during the Deadball Era, pitchers were “allowed to use altered baseballs and trick pitches.” This resulted in better pitching performances and decreased offensive output ([Netshrine, 2009](#)).

The focus of the educational exercise in this article is the transition from the Designated Hitter Era to the Steroids Era (which corresponds to the Power Era as defined above by Schell). Widely considered to have begun in 1993 or 1994, the Steroids Era has been referred to as such because of the perceived widespread use of steroids and other banned, performance-enhancing drugs ([Baseball Prospectus, 2009](#)). The exact proportion of Major League players that used (or currently use) steroids or other performance-enhancing drugs is not publicly released, though a player testing program instituted in 2003 found that between 5-7% of players were users ([Mitchell, 2007](#)). As will be evident later in this article, offensive performance sharply increased during the beginning of the Steroids Era and remained at relatively high levels through the 2008 season. While outside the scope of this educational exercise, a number of researchers have attempted to analyze the nature of the link between steroid use and performance in baseball (e.g. [Schmotzer et al., 2008](#), [Schmotzer et al., 2009](#), and [Albert 2007](#)). Each of these papers presents evidence that suggests that the use of performance-enhancing drugs resulted in improved offensive performance.

A key question for this exercise is: “How does one quantify offensive performance in baseball?” Various measures of offensive output by baseball players are available, but the league batting average of all Major League Baseball players is used in this exercise because of its simplicity and clean results. League batting average (*BA*) is calculated as follows for any single season:

$$\text{League BA} = \frac{\text{Total number of safe hits}}{\text{Total number of at bats}}$$

According to [MLB Official Rules \(2009\)](#), a player does not incur an “at bat” when he: Hits a sacrifice bunt or sacrifice fly, is awarded first base on four called balls, is hit by a pitched ball, or is awarded first base because of interference or obstruction.

Note that some will likely question the use of batting average as an appropriate measure of offensive performance in the context of performance-enhancing drug use. For example, in their studies of the effects of performance-enhancing drugs on offensive performance in baseball, [Schmotzer et al. \(2008\)](#) and [Schmotzer et al. \(2009\)](#) use a “runs created per 27 outs” metric and [Albert \(2007\)](#) uses a “home run rate” metric. However, for pedagogical purposes of ease of use and understanding, league batting average is used in this article as the measure of offensive performance. For those that may prefer an alternative metric, [Appendix B](#) provides data on home runs per game, and discusses its potential use.

[Figure 1](#) is a plot of league batting average data from 1901 to 1992. Note that the vertical lines separate the data by era. As is evident in this figure, both the Deadball and Lively Ball eras (from 1901-1919 and 1920-1946, respectively) featured offensive play fundamentally different from that in the last 40 years. The Post-World War II era, from 1947 to 1962, was a relatively stable period for offensive performance. The Big Strike Zone Era of 1963-1968 was characterized by a reduction in offensive performance due to an expansion of the definition of the strike zone. The motive behind the change was a desire to improve the pace at which games were played. However, this rule change relating to the strike zone was overturned for the 1969 season onward. Following this was the Designated Hitter Era, which was defined by [Schell \(2005\)](#) to last from 1969 to 1992 though, as will be “discovered” by students in this educational exercise, the implementation of the designated hitter rule actually did not occur until 1973. This rule, which is used in the American League, allows for a “designated hitter” to bat in place of the pitcher. Advent of the designated hitter has increased team batting performance in the American League.

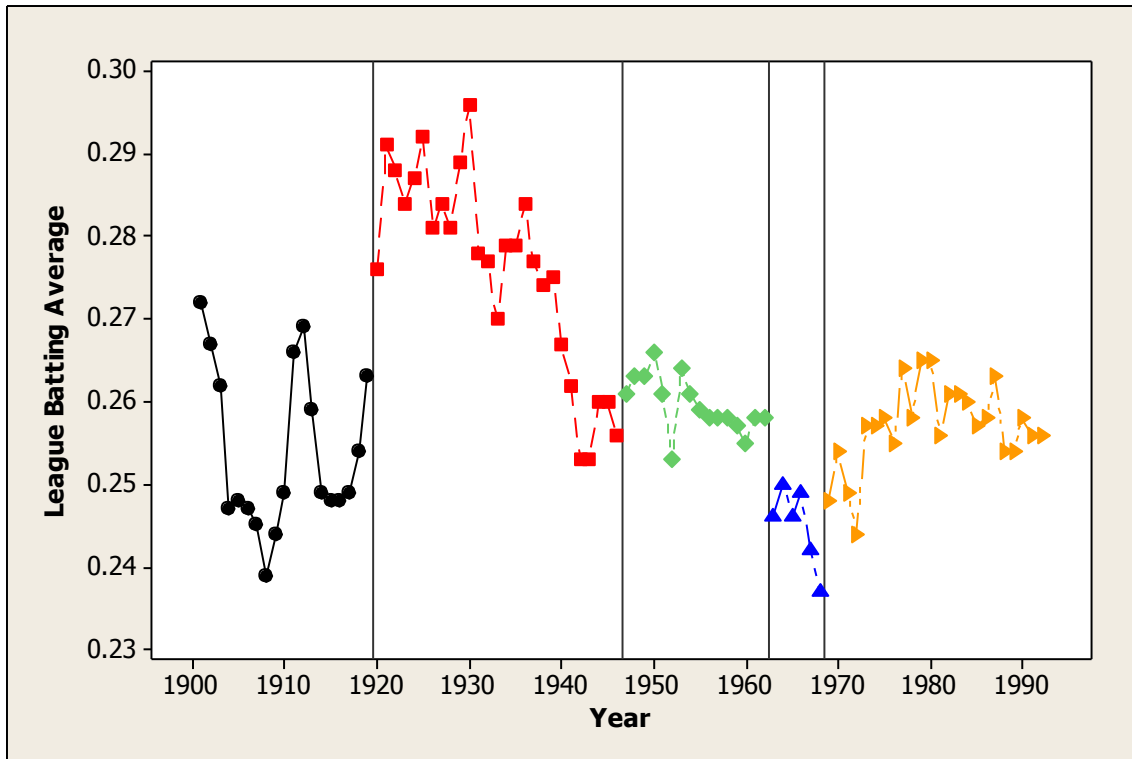


Figure 1: Plot of League Batting Average (1901-1992)

Although league batting average data are available for all eras of modern baseball history, only data from the relatively stable 1969 - 1992 period defined by [Schell \(2005\)](#) are used as the baseline for calculating control chart trial limits in this exercise. These data, together with league batting averages for the 1993 to 2008 seasons, are provided in [Appendix A](#) in tabular form and in Excel format in the attached `battingaverage.xls` file. The data are from the [Baseball Almanac \(2009\)](#) and are to be provided to the students.

3. Statistical Process Control Charts

The primary focus of this exercise is the construction and application of statistical process control charts. These charts were first introduced by William Shewhart in the 1920's and have been used for many years as a graphical method for the monitoring of process behavior and variability in both manufacturing and service contexts ([Montgomery, 2005](#)). Control charts allow for the identification of assignable cause variation, i.e. variation that is not purely random. In practice, efforts are made to identify and address assignable cause variation, thereby improving process performance.

For this exercise, the Individuals (*I*) and Moving Range (*MR*) charts are utilized, as the data are in the form of single observations (the league batting average) for each baseball season. The *I*

and *MR* charts allow for monitoring of central tendency and variation, respectively, of a process and are often used in tandem. These charts are commonly employed in situations where automated inspection technology is utilized and every unit is inspected, where the production rate is slow and it is inconvenient to allow samples of greater than one item to accumulate, or other situations where only one measurement is meaningful or available for a given sample (Montgomery, 2005). If students are learning about other types of control charts, it may be pointed out that a *p*-chart could be constructed for this data as an alternative to *I* and *MR* charts. This is because the data has the basic characteristics of binomial data in that each at-bat may be considered as a trial that results in a hit or not, and consequently each season's batting average is the proportion of at-bats that resulted in a hit.

The initial phase of development of statistical process control charts involves the construction of trial control limits for a baseline period of data. To calculate the limits for *I* and *MR* control charts, the average moving range and average observation values are needed. Moving range, *MR*, provides a measure of process variation and is calculated from two consecutive observations by the following:

$$MR_i = |x_i - x_{i-1}|$$

Students should recognize that a moving range cannot be calculated for the first observation in a data set because there is no prior value from which to calculate a difference. Consequently, an *MR* control chart will have the first time period left blank and begin with the moving range value for the second period.

For the *I* chart, the upper (UCL) and lower (LCL) control limits as well as the centerline (CL) are then calculated by the following:

$$UCL = \bar{X} + 3 \frac{\overline{MR}}{d_2}$$

$$CL = \bar{X}$$

$$LCL = \bar{X} - 3 \frac{\overline{MR}}{d_2}$$

where \bar{X} is the mean of the observations, \overline{MR} is the mean of the moving ranges, and d_2 is a sigma conversion factor related to the sample size of the moving range. For a typical moving range of two observations, $d_2 = 1.128$ (Hines et al., 2004).

For the Moving Range (*MR*) chart, the control limits are calculated by the following:

$$UCL = D_4 \overline{MR}$$

$$CL = \overline{MR}$$

$$LCL = D_3 \overline{MR}$$

where D_3 and D_4 are parameters related to the number of observations in each moving range. For a typical moving range of two observations, $D_3 = 0$ and $D_4 = 3.267$ ([Hines et al., 2004](#)).

4. Exercise Activities and Guidance

An inherent difficulty in teaching students about control charts is the lack of “real” process data that can be used to generate process control charts for in-class demonstration and analysis. Random data can be simulated via Minitab, Excel, or other software packages, but pedagogical research suggests that students are often more engaged in a classroom topic when real data are used, and student involvement is encouraged ([Snee, 1993](#)). Such student engagement is often seen when involvement is encouraged in statistics courses (for examples, see [Bradstreet \(1996\)](#) and [Yilmaz \(1996\)](#)). For these reasons, the exercise given in this article has the advantage of giving students the opportunity to analyze “real world” data. As these data are from a controversial and current issue related to sports in the United States, the interest level of most students for this topic can be expected to be high.

In this exercise, students are to complete the following tasks to construct control charts for the league batting average data for the baseline period of 1969 to 1992. Each of these tasks is described afterward in further detail.

1. Assess the normality of the league batting average data for the baseline period of 1969 - 1992.
2. Construct trial control limits for the 1969-1992 data and identify out-of-control points, if any, during this baseline period.
3. Determine if the out-of-control points can be attributed to assignable cause variation. If so, remove the points and recalculate the trial control limits.
4. With the trial control limits fixed, plot subsequent data from the process (1993-2008) on the control charts and evaluate.

Both the I and MR control charts require that data from the process under study be normally distributed. Before proceeding with control chart development, students should be asked to perform a test for normality on the league batting average data. [Figure 2](#) is a normal probability plot for the league batting average from 1969-1992. Based upon the shape of the data in the plot as well as the corresponding Anderson-Darling normality test p -value ($p = 0.178$), it is reasonable to assume that the league batting average over this period of time is normally distributed.

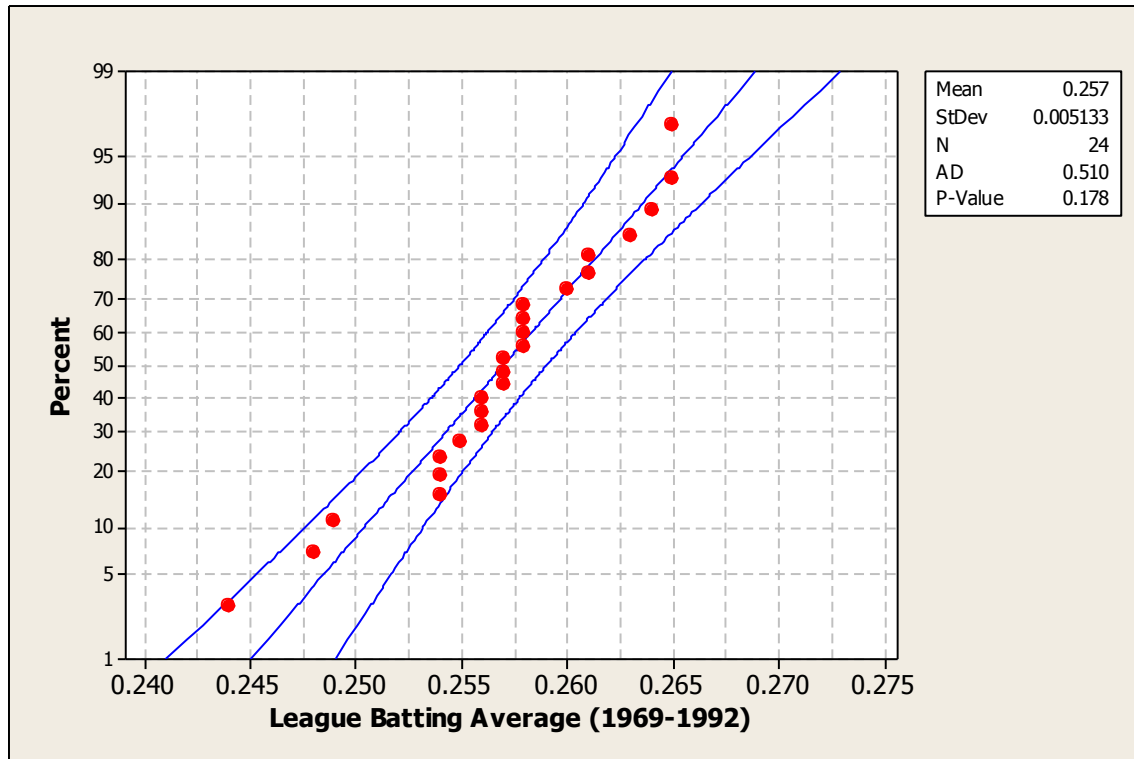


Figure 2: Normal Probability Plot of League Batting Average (1969-1992)

[Figure 3](#) shows I and MR trial control limits generated in the Minitab statistical software package from data over the period of 1969-1992, which is the Designated Hitter Era as defined by [Schell \(2005\)](#). The league batting average over this period appears to be relatively stable and in-control, with the possible exception of the 1969-1972 seasons when the league batting average appears to be somewhat lower and clearly dips to an out-of-control point in 1972. As when dealing with output from a production process, students should be asked to consider whether there is assignable cause variation associated with these points.

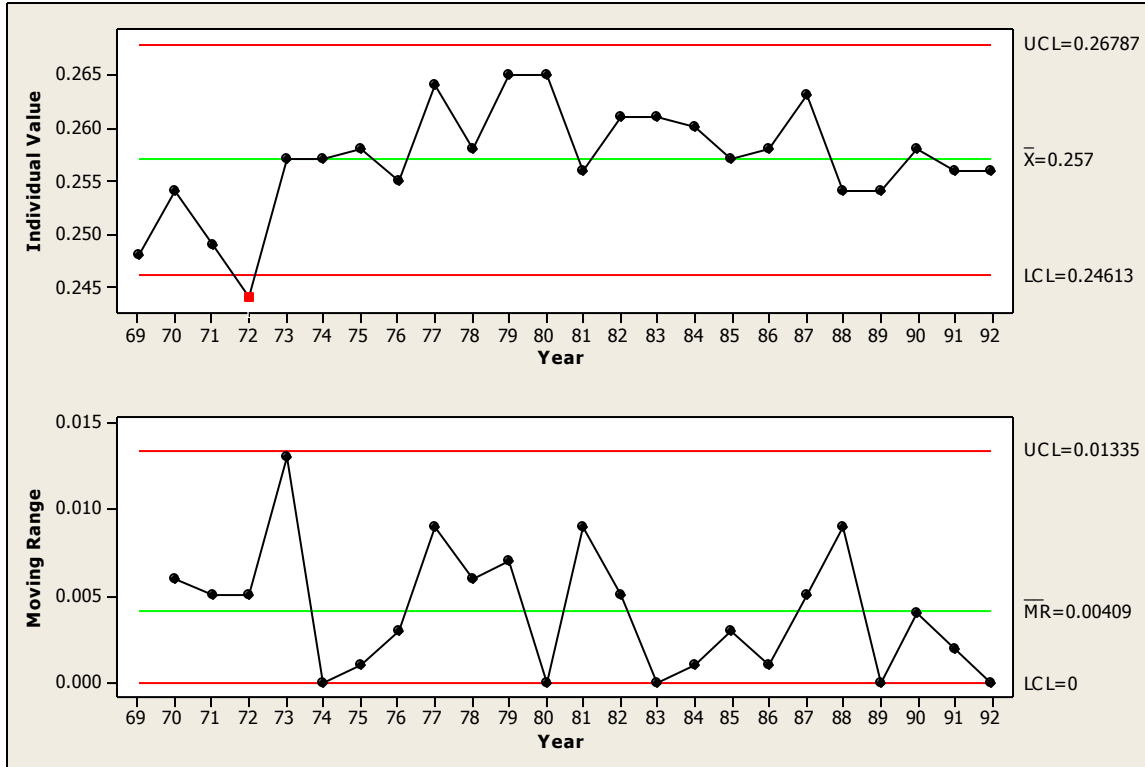


Figure 3: I/MR Charts for League Batting Average (1969-1992)

[Treder \(2005\)](#) addressed the uniqueness of the 1969-1972 seasons with the following points:

- Although Schell refers to the period beginning in 1969 as the Designated Hitter Era, Treder notes that the American League actually did not implement the designated hitter rule until 1973. The effect of this implementation was immediate, as the league batting average rose by 5% in 1973 compared to 1972. Accordingly, the absence of the designated hitter rule in the 1969 to 1972 seasons can be viewed as an assignable cause of variation which may explain why league batting averages were lower during these seasons than in the seasons from 1973 onwards.
- Furthermore, the 1972 season was disrupted at its onset by a players strike. This strike, the first in the 20th century, led to a 13-day delay in starting the season. Pre-season training was affected, and games that were to take place during the strike period were not replayed. This may have been another assignable cause contributing to lower performance in 1972.
- In 1969, rules were changed so that the pitching mound was lowered with the intent of reducing pitchers' effectiveness and improving offensive output. Treder contends that this rule change had a transitory effect in raising offensive performance above that of 1967 or 1968, but that this effect had dissipated by 1972 as "forces favoring pitchers reasserted themselves and overrode the impact of the 1969 rule changes."

Given the above, it can be argued that there is sufficient assignable cause variation associated with the 1969 to 1972 league batting average data to justify removing these data points from the control limit calculations. It can be pointed out to students that this is similar to a real-world business process in the sense that we might think *a priori* that the process was stable during our baseline period of data, but it is necessary to calculate trial control limits in order to examine whether, in fact, the process experienced any non-random, assignable causes of variation. After removing the assignable cause variation of 1969-1972, a normal probability plot of the data for 1973-1992 confirms that it still appears to be normally distributed. Revised trial control limits are calculated next using only the 1973-1992 data as the baseline period, which results in the control charts shown in [Figure 4](#). In this figure, all points appear to be in control and exhibit only random cause variation. Accordingly, we may assume that the control limits in this chart now reflect only random variation and that they can be used to monitor variation in league batting average for subsequent seasons from 1993 onward.

In addition to the normality assumption for *I* and *MR* charts, another issue that may be discussed in an advanced course is the assumption made in standard control charts that samples are independent and identically distributed over time. If samples have a positive autocorrelation, the control chart may experience a higher rate of false alarms than expected. In this exercise, advanced students can be asked to consider whether the batting average data is independent over time. Since a portion of the batters and pitchers who play in a given season will carry-over to the next season, it is possible that there is some degree of dependency between adjacent seasons. Advanced students could be asked to further explore the concepts and methods surrounding this often overlooked issue (refer to [Montgomery \(2005\)](#), for example, for a discussion of SPC with correlated process data).

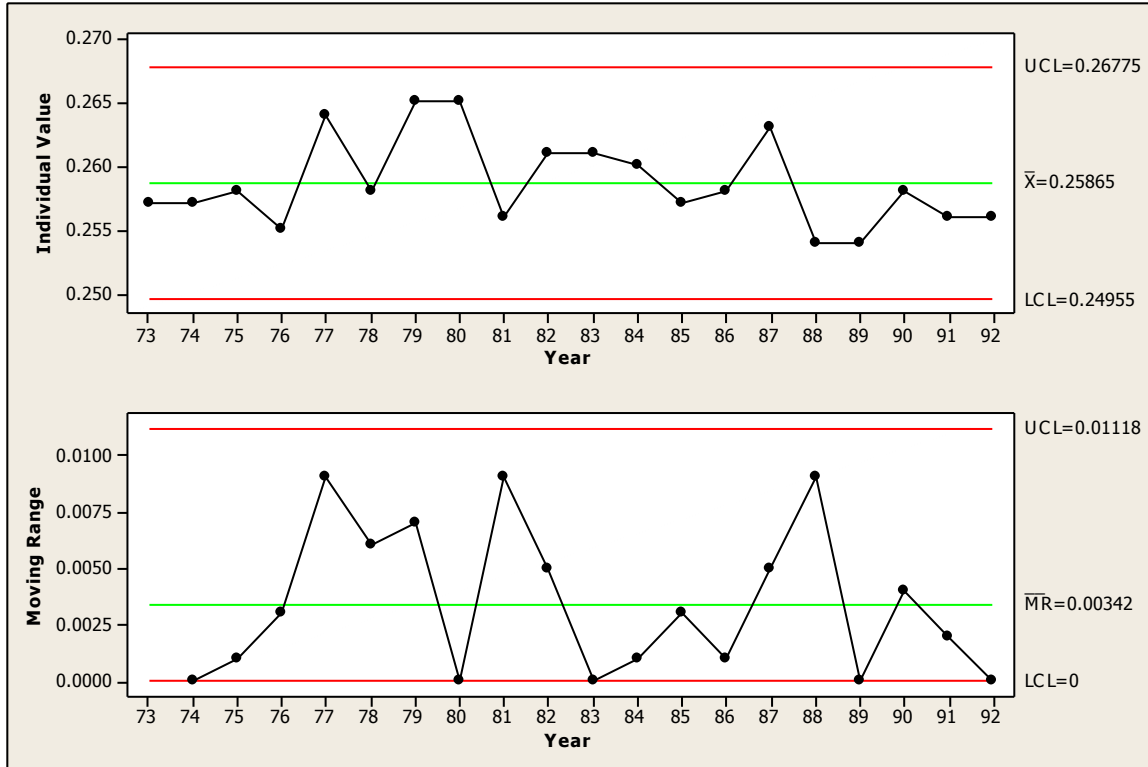


Figure 4: *I*/*MR* Charts for League Batting Average (1973-1992)

Next, *I* and *MR* control charts for the league batting averages for the entire 1973-2008 period are shown in [Figure 5](#), with the control limits fixed as indicated previously in [Figure 4](#). Note that the vertical dashed line delineates the data used for development of the trial control limits for the baseline period (1973-1992), and the subsequent data that is being evaluated in comparison to these control limits (1993-2008). This enables us to consider whether batting averages from 1993 onwards differ significantly from those observed during the 1973-1992 baseline period.

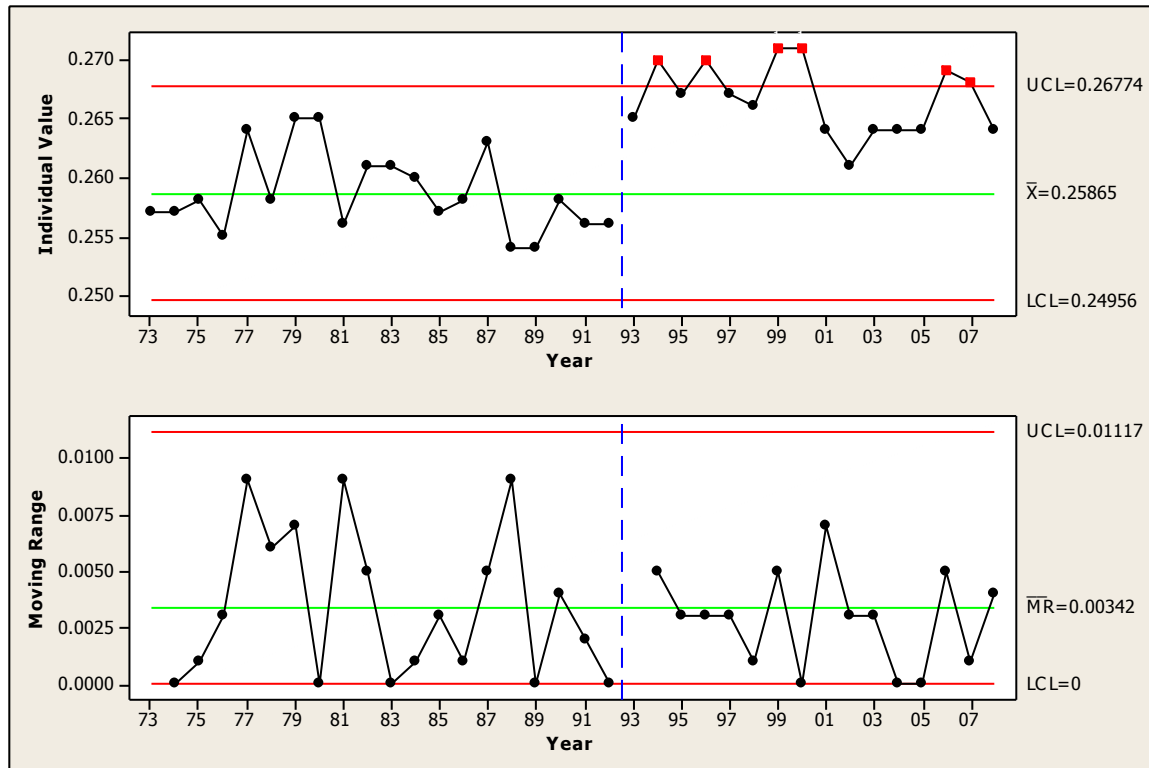


Figure 5: *I*/*MR* Charts for League Batting Average (1973-2008)

While no significant changes are observed in the *MR* chart of [Figure 5](#), the league batting average “process” on the *I* chart is out of control throughout the 1993-2008 period according to several supplemental rules for interpreting control charts. These “runs rules” or tests for interpretation of control charts can be found in [Western Electric Company \(1956\)](#) and [Nelson \(1984; 1985\)](#) and are as follows:

- Six points on the *I* chart are beyond the three-sigma control limits. Any one of these points indicates the likelihood of assignable cause variation.
- All sixteen points from 1993-2008 are above the centerline, which indicates an upward shift in the process. According to [Western Electric \(1956\)](#) and [Montgomery \(2005\)](#), eight (or more) points in a row falling on the same side of the mean are evidence of a process shift from some assignable cause.
- The 1997 and 1998 seasons also exhibit another out-of-control pattern of variation. When examining the three-year period of 1997-1999, these two seasons constitute two (or more) out of three points in a row that fall beyond the two-sigma limit, which indicates non-randomness regardless of the value for the third point (1999). Similarly, the 1993 and 1995 seasons indicate assignable cause variation regardless of the value for the 1994 season, and the same can be said for the 1995 and 1997 seasons regardless of the value for the 1996 season.
- In addition, the 2001, 2003, 2004 and 2005 seasons constitute four (or more) out of five points in a row (2001-2005) that fall beyond the one-sigma limit, which is another pattern indicating the likelihood of assignable cause variation.

It should be pointed out to students that the above runs rules are based on the assumption of independence of the samples. As a learning point, advanced students can be asked to identify what the effect would be of dependence or autocorrelation in the samples, with the correct answer being that it would increase the probability of such runs in the data and thereby increase the chance of false alarms from the control chart. These control charts and their potential implications are discussed further in the following section.

5. Analysis

Having constructed the control charts shown above in [Figure 5](#), students should be asked to analyze them. As explained above, the *I* chart indicates that the process is out of control, and a considerable shift in the league batting average is evident from the 1993 season onward. Thus, one or more changes likely occurred that resulted in this change in process behavior. Had statistical process control charting techniques been used to monitor the league batting average, this change would have been readily noticed. The task now at hand is to investigate the cause(s) behind this change. Can the use of banned, performance-enhancing drugs be solely to blame?

Although the Steroids Era is widely thought to have begun in or around 1993, which coincides with this increase in batting performance, several other notable events occurred in Major League Baseball that may have contributed to the increase in offensive performance:

- While Major League Baseball's definition of a "strike" did not change during the 1992-2008 period, a trend emerged where "inside" pitches (pitches close to the batter, but still passing over home plate) were not being called strikes by the umpires. This led to a *de facto* decrease in the size of the strike zone, which was a development favorable to the offense ([Rader and Winkle, 2008](#)). In 2001, Major League Baseball responded to this development by decreeing that umpires should follow the rulebook definition of a strike ([Walker, 2001](#)). As seen in [Figure 5](#), there was a notable decline in league batting average from the 2000 to 2001 seasons. By 2006, however, offenses seemed to have adjusted to the newly "expanded" strike zone, as the league batting average returned to a higher level.
- Several teams moved to new stadiums that are considered to be more favorable to offensive production. For an example of this, see the "Park Factors" statistic ([MLB Park Factors, 2009](#)). This statistic is calculated for each of the 30 existing MLB stadiums and indicates whether a stadium is favorable to either the offense or the defense. Of the 16 stadiums that favor the offense (from 2008 season data), eleven were opened for play in the 1992-2008 timeframe.
- Major League Baseball added two new teams in 1993: the Colorado Rockies and the Florida Marlins. The expansion to Colorado was notable, as the Rockies play at a high altitude in Denver. There is increased offensive performance at high altitudes due to increases in ball flight distances ([Bahill et al., 2009](#)).
- There has been speculation that the physical construction of the baseballs used in Major League Baseball since the 1990s has been different than that during previous periods. Allegedly, baseballs were "juiced", or constructed in such a way as to produce more offense, especially home runs. Opinions on this, however, are widely divided. [Sloat \(2000\)](#), for example, outlines several arguments and studies supporting the assertion that

baseballs were modified during this period, while [Rader and Winkle \(2002\)](#) contend that no definitive evidence has been produced.

- It is conceivable that the increase in batting average is due to a cohort of exceptionally strong batters and/or weak pitchers that happened to play during the time period of 1993 to 2008. This correlates with the addition of two, new MLB teams in 1993, as some have suggested that expansion “dilutes” pitching talent and provides an opportunity for improved offensive performance ([Rader and Winkle, 2002](#)). However, when MLB expanded again in 1998 with the addition of two more teams, there was no significant change to league batting average.

In summary, performance-enhancing drugs likely were a factor in the increase in offensive performance during the Steroids Era. However, the relative degree to which offensive performance was affected by drug use is not readily apparent. Students should be asked to discuss this. In particular, a learning point can be made regarding the use of statistical process control charts: These charts have the capability to identify when a significant process change has occurred, but cannot identify the cause. They serve only as a warning device, and further investigation is required to determine the specific cause of the process change and to decide on what corrective action would be appropriate.

6. Conclusions

In this article, a statistical process control charting exercise is proposed that allows students to examine the behavior of offensive performance in Major League Baseball before and during the so-called Steroids Era of baseball history. Students are given an opportunity to construct trial control limits, identify random and assignable cause variation, and analyze completed process control charts. The exercise utilizes real-world data related to a controversial and timely topic in American sports. Therefore, student interest and involvement in the exercise should be high.

Appendix A: Data for League Batting Average (1969-1992 and 1993-2008)

An Excel file containing this data set is available for download from JSE at:

<http://www.amstat.org/publications/jse/v19n1/battingaverage.xls>

Year	League BA	Year	League BA
1969	0.248	1993	0.265
1970	0.254	1994	0.270
1971	0.249	1995	0.267
1972	0.244	1996	0.270
1973	0.257	1997	0.267
1974	0.257	1998	0.266
1975	0.258	1999	0.271
1976	0.255	2000	0.271
1977	0.264	2001	0.264
1978	0.258	2002	0.261
1979	0.265	2003	0.264
1980	0.265	2004	0.264
1981	0.256	2005	0.264
1982	0.261	2006	0.269
1983	0.261	2007	0.268
1984	0.260	2008	0.264
1985	0.257		
1986	0.258		
1987	0.263		
1988	0.254		
1989	0.254		
1990	0.258		
1991	0.256		
1992	0.256		

Appendix B: Data and Analysis of Home Runs per Game (1969-1992 and 1993-2008)

As an alternative metric of offensive performance, data on home runs per game, [Baseball-Reference.com \(2009\)](#) are provided at the end of Appendix B and in Excel format in the attached homeruns.xls file. This metric may be appealing because of the intuitive connection between steroid use, strength, and hitting home runs. However, the results are not quite as clean as those for the league batting average data shown in the main body of this paper. Consequently, the batting average data is recommended for purposes of the educational exercise and this data on home runs per game is provided as an additional reference.

Construction of control charts for the home runs per game data follows the same procedures as used for the batting average data. Using 1969-1992 data as the baseline period for the trial control limits, the 1969-1972 seasons do not appear out of control (unlike with the batting average data) even though the designated hitter rule was not actually implemented until 1973. This result implies that the designated hitter rule is not an assignable cause impacting home runs per game and therefore it is not necessary to remove these seasons and revise the trial control limits. However, the control chart for the baseline period does reveal that the value for the 1987 season exceeds the upper control limit and is in the same range of values as the post-1993 seasons of the Steroids Era. The 1987 season's unusually high value has been speculated to be due to no apparent cause or to be the result of "juicing" the ball ([Deford, 1987](#)), and consequently this outlier result is a distraction that clouds the exercise and its motivating purpose for students of examining the transition into the Steroids Era. As to whether or not the 1987 season should be removed from the calculation of the control limits, the literature does not appear to have a consensus regarding an out of control point during the baseline period for which no assignable cause is found. [Montgomery \(2005\)](#) suggests that it is permissible to go either way: remove the point as though an assignable cause has been found and revise the trial control limits, or retain the point and trial control limits as they are.

When data for the 1993-2008 seasons from the Steroids Era are added to the baseline control chart, it can be seen that there is indeed a pronounced shift upwards, similar to that seen with the batting average data. In sum, the home runs per game metric of offensive performance gives similar results to the batting average data, though it is not as amenable to the purposes of this educational exercise.

An Excel file containing this data set is available for download from JSE at:

<http://www.amstat.org/publications/jse/v19n1/homeruns.xls>

<u>Year</u>	<u>HR/Game</u>	<u>Year</u>	<u>HR/Game</u>
1969	0.80	1993	0.89
1970	0.88	1994	1.03
1971	0.74	1995	1.01
1972	0.68	1996	1.09
1973	0.80	1997	1.02
1974	0.68	1998	1.04
1975	0.70	1999	1.14
1976	0.58	2000	1.17
1977	0.87	2001	1.12
1978	0.70	2002	1.04
1979	0.82	2003	1.07
1980	0.73	2004	1.12
1981	0.64	2005	1.03
1982	0.80	2006	1.11
1983	0.78	2007	1.02
1984	0.77	2008	1.00
1985	0.86		
1986	0.91		
1987	1.06		
1988	0.76		
1989	0.73		
1990	0.79		
1991	0.80		
1992	0.72		

References

- Albert, J. (2007). "Numb3rs, Sabermetrics, Joe Jackson, and Steroids", *STATS*, 48, 3-6.
- Bahill, A., Baldwin, D. and Ramberg, J. (2009). "Effects of Altitude and Atmospheric Conditions on the Flight of a Baseball", *International Journal of Sports Science and Engineering*, 3, 2, 109-128.
- Baseball Almanac (2009). <http://www.baseball-almanac.com/>, Last accessed October 2009.
- Baseball Prospectus – Baseball Between the Numbers (2009). <http://www.baseballprospectus.com/article.php?articleid=4845>, Last accessed October 2009.
- Baseball-Reference.com (2009). <http://www.baseball-reference.com/leagues/MLB/bat.shtml>, Last accessed October 2010.
- Bradstreet, T.E. (1996). "Teaching Introductory Statistics So That Non-Statisticians Experience Statistical Reasoning", *The American Statistician*, 50, 1, 69-78.
- Deford, F. (1987). "Rabbit Ball: Whodunit?", *Sports Illustrated*, July 27, p. 94.
- Hines, W., Montgomery, D., Goldsman, D., and Borror, C. (2004). *Probability and Statistics in Engineering*, 4th Edition, Wiley & Sons, Inc., Hoboken, NJ.
- Mitchell, G.J. (2007). "Report to the Commissioner of Baseball of an Independent Investigation into the Illegal Use of Steroids and Other Performance Enhancing Substances by Players in Major League Baseball", <http://files.mlb.com/mitchrpt.pdf>, Last accessed October 2009.
- MLB Official Rules (2009). http://mlb.mlb.com/mlb/official_info/official_rules/official_scorer_10.jsp, Last accessed October 2009.
- MLB Park Factors (2009). http://espn.go.com/mlb/stats/parkfactor/_/year/2008, Last accessed October 2009.
- Montgomery, D. (2005). *Introduction to Statistical Quality Control*, 5th Edition, Wiley & Sons, Inc., Hoboken, NJ.
- Nelson, Lloyd S. (1984). "The Shewhart Control Chart—Tests for Special Causes," *Journal of Quality Technology*, 16(4), 237 – 239.
- Nelson, Lloyd S. (1985). "Interpreting Shewhart Control Charts," *Journal of Quality Technology*, 17(2), 114 – 116.
- Netshrine – Baseball Eras Defined (2009). <http://www.netshrine.com/era.html>, Last accessed October 2009.

Rader, B. and Winkle, K. (2002). "Baseball's Great Hitting Barrage of the 1990s", *NINE: A Journal of Baseball History and Culture*, 10, 2, 1-17.

Rader, B. and Winkle, K. (2008). "Baseball's Great Hitting Barrage of the 1990s (and Beyond) Reexamined", *NINE: A Journal of Baseball History and Culture*, 17, 1, 71-96.

Schell, M.J. (2005). *Baseball's All-Time Best Sluggers: Adjusted Batting Performance from Strikeouts to Home Runs*, Princeton University Press, Princeton, NJ.

Schmotzer, B., Kilgo, P.D. and Switchenko, P. (2009). "'The Natural' The Effect of Steroids on Offensive Performance in Baseball", *Chance*, 22, 2, 21-32.

Schmotzer, B., Switchenko, B., and Kilgo, P. (2008). "Did Steroid Use Enhance the Performance of the Mitchell Batters? The Effect of Alleged Performance Enhancing Drug Use on Offensive Performance from 1995 to 2007", *Journal of Quantitative Analysis in Sports*, 4, 3, 1-14.

Sloat, B. (2000). "They're lively to the core, tests reveal; Why today's baseballs go farther", *The Plain Dealer* (Cleveland), September 29, 2000.

Snee, R. (1993). "What's Missing in Statistical Education?", *The American Statistician*, 47, 149-154.

Treder, S. (2005). "1972: The Year That Changed Everything", *NINE: A Journal of Baseball History and Culture*, 13, 2, 1-18.

Walker, B. (2001). "New Strike Zone Makes Its Debut", *AP Online*, March 2.

Western Electric Company (1956), *Statistical Quality Control Handbook*. (1st ed.), Indianapolis, IN: Western Electric Co.

Yilmaz, M.R. (1996). "The Challenge of Teaching Statistics to Non-Specialists", *Journal of Statistics Education*, 4, 1, <http://www.amstat.org/publications/jse/v4n1/yilmaz.html>.

Stephen E. Hill, Ph.D.
Weber State University
John B. Goddard School of Business & Economics
Department of Business Administration
3802 University Circle, Ogden, UT 84408-3802
Email: stephenhill@weber.edu
Phone: 801-626-6075

Shane J. Schvaneveldt, Ph.D.
Weber State University
John B. Goddard School of Business & Economics
Department of Business Administration
3802 University Circle, Ogden, UT 84408-3802
Email: schvaneveldt@weber.edu
Phone: 801-626-6075

[Volume 19 \(2011\)](#) | [Archive](#) | [Index](#) | [Data Archive](#) | [Resources](#) | [Editorial Board](#) |
[Guidelines for Authors](#) | [Guidelines for Data Contributors](#) | [Guidelines for Readers/Data](#)
[Users](#) | [Home Page](#) | [Contact JSE](#) | [ASA Publications](#)