STATISTICAL PRACTICE

This department publishes articles of interest to statistical practitioners. Innovative applications of known methodology may be suitable, but sizable case studies should be submitted to other journals. Brief descriptions and illustrations of new developments that are potentially useful in statistical practice are appropriate. Acceptable articles should appeal to a substantial number of practitioners.

Modeling Pitcher Performance and the Distribution of Runs per Inning in Major League Baseball

Bernard ROSNER, Frederick MOSTELLER, and Cleo YOUTZ

The distribution of the number of batters faced and the number of runs scored in an inning is fundamental to modeling pitcher performance in major league baseball. Until 1984 sufficiently detailed data on a play-by-play basis were available only for special games, for example, World Series games. Play-by-play data have become available for each league beginning with the 1984 season, and we use them for the 1990 American League season. We found that a minor change from the negative binomial model provided an adequate fit to the distribution of the number of batters faced. However, the distribution of the number of runs scored is complex and involves a convolution of the distribution of the number of batters faced and the conditional distribution of the number of runs scored given a specific number of batters faced. We modeled the latter by a truncated binomial distribution where the probability of a run scoring was a logistic function of the number of batters faced. The goodness of the fit of each of the models was tested on 77 starting pitchers who pitched in the first three innings of enough games to total at least 30 innings, all cumulating for these pitchers to 5,639 innings during the 1990 American League season.

KEY WORDS: Baseball modeling; Negative binomial distribution; Sports statistics.

This article develops a model for the distribution of the number of batters faced and the number of runs scored against starting pitchers in the early innings of American League baseball games based on data from 1990. The simplest models, such as the negative binomial distribution for the number of batters faced with fixed probabilities of getting a batter out, had to be adjusted. By attending to the most-used starting pitchers and restricting the analysis to the first three innings, we can defer dealing with the effects of tiring and of the introduction of relief pitchers for later research.

The rationale for considering such models is to introduce a stochastic element in the assessment of pitcher performance. Most traditional baseball statistics [such as the earned run average (ERA)] do not allow for random variation in assessing performance. Thus one can rank pitchers according to ERA, but one cannot translate such rankings into assessing the probability that pitcher A will allow fewer runs than pitcher B on any given day. Such stochastic paired comparisons offer a richer set of consequences in their descriptions than the usual point estimate that assesses pitcher performance.

1. MODEL

For an inning under consideration let

x = number of runs scored
$h(x)$ = probability of exactly x runs being scored
b = number of men left on base, $0 \leq b \leq 3$
N = number of batters the pitcher faced
$f(N)$ = probability of facing exactly N batters in the inning
$g(x|N)$ = conditional probability of x runs given N batters.

We ignore unfinished innings in games interrupted by such matters as darkness, weather, acts of God, or local laws except as discussed below. The probability of scoring exactly x runs in an inning is given by

$$h(x) = \sum_{N=x+3}^{x+6} f(N)g(x|N). \tag{1}$$

The reason for the constraint on N in the summation is that by the end of an inning, a batter must either score, be left on base, or be put out. (A player who comes to the plate but does not meet one of these three conditions is not counted as a batter. Replacement batters or runners are regarded as indistinguishable from the original batter or runner.) The number of batters is therefore $N = x + b + 3$, and so for a

Bernard Rosner is Professor of Medicine (Biostatistics), Harvard Medical School, Boston, MA 02115. Frederick Mosteller is Roger I. Lee Professor of Mathematical Statistics, Emeritus, Departments of Statistics and Health Policy and Management, Harvard University, Cambridge, MA 02138. Cleo Youtz is Mathematical Assistant, Department of Statistics, Harvard University, Cambridge, MA 02138.

given value of x the smallest and largest numbers of batters are $x+3$ and $x+6$, respectively.

For parsimony and interpretability we want to specify a parametric form for f and g, and thus reduce the number of parameters needed to fit a model. For this purpose let

$$p = \text{probability of an out}$$

(where, without loss of generality, we assume that the last batter faced is an out). Because the length of an inning is determined mainly (although not entirely) by "at bats" that are sorted into "outs" and "not outs," and because three outs end an inning, the negative binomial distribution is a natural distribution to explore because its main feature is trials, here at bats, until a predetermined number of successes occurs. Therefore, we consider a negative binomial model for f, namely

$$f(N) = \binom{N-1}{2} p^3 (1-p)^{N-3}, \qquad N \geq 3. \tag{2}$$

Upon examining actual data we found that model (2) underestimated the probability of facing exactly three batters and overestimated the probability of facing exactly four batters, with other outcomes being closely predicted. These deviations from the negative binomial may be primarily due to the occurrence of double plays, where two outs occur with a single at-bat, or where a runner is caught stealing second after successfully reaching first base. Therefore, we modified the model in (2) as follows:

$$f_1(N) = \begin{cases} p^3 + \lambda & \text{if } N = 3 \\ 3p^3(1-p) - \lambda & \text{if } N = 4 \\ f(N) & \text{if } N \geq 5, \end{cases} \tag{3}$$

where λ (usually positive for a given pitcher) is a rough indicator of the propensity of a pitcher and his team to either achieve double plays and/or to prevent a runner on first from obtaining a big lead. Appendix A shows that the MLE's of $f_1(3)$ and $f_1(4)$ must always lie between 0 and 1.

To model $g(x|N)$, the conditional distribution of the number of runs scored (x) given the number of batters faced (N), we let p_1 = probability that a batter scores a run given that he is a nonout by the end of the inning. Under the simplest model we let x be binomially distributed with parameters $N-3$ and p_1. However, x is constrained because the number of men on base at the end of the inning, b or equivalently $(N-3-x)$, must satisfy the inequality $0 \leq N-3-x \leq 3$, which is equivalent to $N-6 \leq x \leq N-3$. Therefore, we model x by a truncated binomial distribution of the form

$$g(x|N) = \frac{\binom{N-3}{x} p_1^x (1-p_1)^{N-3-x}}{\sum_{y=N_1}^{N-3} \binom{N-3}{y} p_1^y (1-p_1)^{N-3-y}} \quad \text{if } N > 3,$$

where

$$x = N_1, \ldots, N-3$$

and

$$N_1 = \max(0, N-6)$$

$$g(0|N) = 1 \quad \text{if } N = 3. \tag{4}$$

An assumption under the model in (4) is that the probability, p_1 (that a batter scores a run given that he is a nonout), is independent of N. However, when actual data were compared to the results of this formulation, we found that this was an unrealistic assumption because p_1 generally increased with increasing N. To relax this assumption we introduce

$$\text{logit } [p_1(N)] = \alpha + \beta(N-3), \tag{5}$$

where α and β are constants fitted for each pitcher.

Then we modified Equation (4) by replacing g by g_1 where $p_1(N)$ replaces p_1, and we get

$$g_1(x|N) = \frac{\binom{N-3}{x} [p_1(N)]^x [1 - p_1(N)]^{N-3-x}}{\sum_{y=N_1}^{N-3} \binom{N-3}{y} [p_1(N)]^y [1 - p_1(N)]^{N-3-y}} \quad \text{if } N > 3,$$

where

$$x = N_1, \ldots, N-3$$

and

$$N_1 = \max(0, N-6)$$

$$g_1(0|N) = 1 \quad \text{if } N = 3. \tag{6}$$

Upon combining Equations (3), (5), and (6), we obtain the overall likelihood for an individual inning given by

$$L(x, N | p, \lambda, \alpha, \beta) = f_1(N) g_1(x|N). \tag{7}$$

2. ESTIMATION

We maximize the likelihood separately for each starting pitcher, using the first three innings of all games started during the 1990 season. To maximize the likelihood we note that f_1 is a function of p and λ, while g_1 is a function of α and β. Therefore, we can maximize each component of the likelihood separately. Upon maximizing f_1 we obtain closed-form expressions for the maximum likelihood estimates (MLE's) of p and λ given by

$$\hat{p} = 1 - \Delta$$

$$\hat{\lambda} = \frac{3\hat{p}^3(1-\hat{p}) I_3 - \hat{p}^3 I_4}{I_3 + I_4}, \tag{8}$$

where

$$\Delta = \frac{-c_1 + c_2 + 2c_3 + \sqrt{(c_1 - c_2 - 2c_3)^2 + 4c_3(3c_1 + c_2 + 3c_3)}}{2(3c_1 + c_2 + 3c_3)}$$

$$c_1 = 3 \sum_{N=3}^{\infty} I_N, \qquad c_2 = 3(I_3 + I_4), \qquad c_3 = \sum_{N=5}^{\infty} (N-3) I_N,$$

and

I_N = number of innings in which the number of batters faced is N.

Table 1. Summary Statistics for Parameter Estimates in Appendix B

Parameter	Mean	SD	N	Range
p	.708	.029	77	(.634, .774)
λ	.024	.052	77	(−.112, .153)
α	−2.261	.692	77	(−3.748, −.335)
β	.638	.236	77	(.157, 1.410)
OBRA	4.183	1.065	77	(2.000, 6.692)
EXRA	4.131	1.024	77	(1.912, 6.649)

Correlation matrix for parameter estimates in Appendix B (N = 77)

	p	λ	α	β	OBRA	EXRA
p	1.0	−.312 (p = .006)	−.164 (p = .15)	.204 (p = .076)	−.791 (p < .001)	−.764 (p < .001)
λ	—	1.0	.237 (p = .038)	−.196 (p = .088)	.230 (p = .044)	.253 (p = .027)
α	—	—	1.0	−.873 (p < .001)	.319 (p = .005)	.387 (p < .001)
β	—	—	—	1.0	−.135 (p = .24)	−.108 (p = .35)
OBRA	—	—	—	—	1.0	.906 (p < .001)
EXRA						1.0

Appendix A gives a sketch of the derivation of Equations (8). To maximize g_1 we use the logistic regression program SAS PROC LOGISTIC to obtain maximum likelihood estimates of α and β, where the database consists of all innings (among the first three) having $N \geq 4$.

Using the maximum likelihood estimates of p, λ, α, and β we obtain an estimate of $h(x)$, that is, of the distribution of the number of runs scored in an inning [i.e., $h(x)$] based on Equation (1). For assessing goodness of fit we also compute the corresponding

$$\text{expected number of innings with } x \text{ runs scored} = nh(x),$$

where n is the number of innings pitched by a specific pitcher.

3. RESULTS

We applied the models in Section 1 to data obtained from starting pitchers in the American League in 1990. The data source is a publicly available magnetic tape provided by Project Scoresheet (1990) consisting of a play-by-play account of each game in the American League during 1990. Data tapes with play-by-play data are available for each league and year for the period 1984–1994. It is almost true that every event that occurs in a game is recorded. A separate record is provided for each at-bat, as well as for selected occurrences during an at-bat (e.g., stealing a base, replacement of a pitcher, etc.). A total of 89,562 records (each involving 75 variables available in ASCII format) are available for games during the 1990 season.

From these data we abstracted for each starting pitcher, for each of the first three innings of each game started, (1) the number of batters faced, (2) the number of outs, and (3) the number of runs scored while this pitcher was in the game. We based our analysis on starting pitchers who pitched at least 30 innings over the season during the first 3 innings of a game (approximately 10 games started). There were 77 American League pitchers who satisfied this criterion during the 1990 season. One complication in the model fitting was the presence of 115 incomplete innings (out of a total of 5,639 innings), where the starting pitcher did not finish the inning. For the purpose of estimation, if a pitcher allowed x runs and faced N batters during an incomplete inning, we made the assumption that if he had finished the inning, the actual number of runs allowed would be equal to x and the actual number of batters faced would be equal to the minimum possible number of batters $= \max(N, x+3)$. We fit the parameters p and λ in Equation (8) for each of the 77 pitchers and present results in Appendix B. Summary statistics for the parameter estimates in Appendix B appear in Table 1.

Table 2 gives, for pitcher Jim Abbott, an example of the computation of $f_1(N)$ and $h(x)$. These are computed from Abbott's empirical joint frequency distribution of N and x. By summing over the corresponding marginal distributions for all 77 pitchers we get the observed marginal distribution of N shown in Table 3, which has been truncated at $N = 10$. We then computed the predicted frequency distribution of the number of batters faced (N) by each pitcher based on Equation (3) (shown in parentheses in Table 2), summed the results over all pitchers, and compared them with the observed distribution of N shown in Table 3.

We assess goodness of fit in Table 3 by the usual chi-squared statistic. By pooling category 9 and categories 10 and beyond we have

$$\chi^2 = \sum_{i=3}^{9} (O_i - E_i)^2/E_i = 4.69 \sim \chi_4^2$$

with significance level = .32. If we pool only the categories from 10 on, the final category alone contributes 7.17 for a total of 10.6 with 5 degrees of freedom, significance level = .06. The model fits the data well, with the exception of the total of the categories at and beyond 10. In this tail

Table 2. *Joint Frequency Distribution and Marginal Probability Distributions*[a] *of Number of Batters Faced (N) and Number of Runs Scored (x) for a Specific Pitcher (Jim Abbott)*

Runs scored (x)	Number of batters faced (N)								Observed (expected)[c] number of innings with x runs scored	Probability of x runs scored (h(x))
	3	4	5	6	7	8	9	10		
0	34[b] (36.7)	18 (18.0)	9 (8.1)	2 (1.9)	—	—	—	—	63 (64.7)	.661
1	—	5 (6.8)	13 (8.5)	4 (4.2)	2 (1.2)	—	—	—	24 (20.7)	.211
2	—	—	2 (2.2)	4 (3.0)	0 (1.9)	0 (.6)	— —	— —	6 (7.6)	.078
3	—	—	—	2 (.7)	1 (1.2)	0 (.8)	1 (.2)	—	4 (3.0)	.030
4	—	—	—	—	0 (.3)	1 (.5)	0 (.3)	0 (.1)	1 (1.2)	.013
5	—	—	—	—	—	0 (.1)	0 (.2)	0 (.1)	0 (.5)	.005
6	—	—	—	—	—	—	0 (.1)	0 (.1)	0 (.2)	.002
7	—	—	—	—	—	—	—	0 (.0)	0 (.0)	.000
Observed (expected)[c] number of innings with N batters faced	34 (36.7)	23 (24.8)	24 (18.8)	12 (9.8)	3 (4.6)	1 (2.0)	1 (.9)	0 (.3)	98	
Probability of N batters faced ($f_1(N)$)	.375	.254	.192	.100	.047	.021	.009	.004		

[a] The probability distribution for N has been truncated at 10 for all pitchers.

[b] Observed (expected) number of innings with N batters faced and x runs scored.

[c] Sums may not agree with totals because of rounding.

the model overpredicts the number of innings compared to the number observed. This may, for example, be due to the tendency of managers to remove a pitcher once the entire opposing lineup has batted around in the inning.

Similarly, by summing the observed and the expected numbers of innings with x runs scored over the tables corresponding to Table 2 for all 77 pitchers, we get the marginal distribution of x shown in Table 4.

Table 3. *Observed and Expected Distribution of N (Based on 77 Starting Pitchers in the American League Season)*

	Number of innings in which pitchers faced exactly N batters			
	Observed		Expected	
N	(O_i)	(%)	(E_i)	(%)
3	2,150	(38)	2,165.3	(38)
4	1,601	(28)	1,610.5	(29)
5	1,019	(18)	1,012.1	(18)
6	488	(9)	492.7	(9)
7	238	(4)	217.7	(4)
8	101	(2)	90.5	(2)
9	38	(1)	36.2	(1)
10+	4	(0)	14.0	(0)
Total	5,639	(100)	5,639	(101)

In assessing degree of fit more than one point of view may be relevant, such as (1) the statistical one with p values for degree of fit, or (2) the absolute size of error when the modeling may not represent the process closely. With large sample sizes p values are likely to indicate that an approximate model is incorrect, as occurs here. But we are also very pleased with the close correspondence between the observed and expected run distributions in Tables 2 and 4, where the percentage comparisons are very close, satisfying attitude (2). However, the goodness of fit was not ideal (chi-square goodness-of-fit statistic $= 21.62 \sim \chi^2_3$, p value less than .0001 due to an underprediction in the number of innings with two or three runs allowed and an overprediction in the number of innings with five or more runs allowed; thus the statistical attitude (1) is dissatisfied).

The traditional approach to assessing the performance of the pitcher is provided by the earned run average. Recall that we restricted our calculations to the first three innings. To relate the parameters fitted in our model to the traditional measure we define the expected run average (EXRA) as $9E[x]$ and compare it for individual pitchers with the

$$\text{observed run average per 9 innings} = \text{OBRA}$$
$$= \frac{9 \times \text{ total runs allowed in first three innings}}{\text{total innings pitched in first three innings}}$$

Table 4. Observed and Expected Distribution of Number of Runs Scored in an Individual Inning (x) (Based on 77 Starting Pitchers in the 1990 American League Season)

Runs scored (x)	Observed number of innings	(%)	Expected number of innings	(%)
0	4,110	(73)	4,139.9	(73)
1	903	(16)	920.1	(16)
2	352	(6)	319.9	(6)
3	172	(3)	137.5	(2)
4	65	(1)	66.9	(1)
5	29	(1)	33.3	(1)
6	5	(0)	15.6	(0)
7	3	(0)	5.9	(0)
Total	5,639		5,639	

(see Appendix B). We make no attempt to distinguish between "earned" and "unearned" runs in the traditional sense because this was unavailable in the database.

A weighted linear regression of observed run average (OBRA) on EXRA was calculated without a constant term and with weights equal to the number of innings pitched, with results given in Table 5.

The regression coefficient is not significantly different from 1, but highly significantly different from 0. A scatterplot and studentized residual plot are shown in Figures 1 and 2. No deviation from linearity, outliers, or heteroscedasticity are apparent from the plots.

Another useful summary is provided by grouping the pitchers by EXRA in one unit increments and computing the mean OBRA within each group. This is given in Table 6.

The correspondence between mean OBRA and EXRA is good with maximum absolute differences of $\leq .13$ in individual groups.

4. DISCUSSION

In this paper we have developed a model for the distribution of runs scored in an individual inning for starting pitchers in the American League in 1990. One use for such a model is to rank pitchers. One index that can be used for this purpose is the expected number of runs (EXRA) allowed per nine innings based on the model. The correlation coefficient between EXRA and the observed number of runs per nine innings (OBRA) is high ($R = .906$). Thus it is not clear that modeling is of much additional value in ranking pitchers. However, modeling does allow one to introduce the concept of random variation to pitcher performance, and allows one to estimate the probability that a specific pitcher (pitcher A) will allow fewer runs than either another pitcher (pitcher B) or than an average pitcher in an individual game, which is not obtainable from traditional baseball statistics such as the ERA. If such models were enhanced by parameters of hitting performance, then the precision of such estimates would improve as well.

Table 5. Weighted Linear Regression of OBRA on EXRA

Variable	Regression coefficient	s.e.	t	p value
EXRA	1.010	.012	84.2	<.001

Table 6. Mean* OBRA Among Subgroups of Pitchers With Given EXRA

EXRA	N	Mean OBRA	Mean EXRA
<3.00	13	2.65	2.55
3.00–3.99	24	3.79	3.66
4.00–4.99	25	4.60	4.55
5.00+	15	5.36	5.48

* Weighted mean where weight = number of innings pitched by individual pitchers.

The models in this paper provide a description of the distribution of the number of runs scored in an individual inning based on four interpretable parameters, p, λ, α, and β. The parameter p is a rough indicator of the probability of a pitcher achieving an out for an individual batter. Parameter estimates range from a high of .774 for Ben McDonald (Baltimore) to a low of .642 for Walt Terrell (Detroit). Based on the *Sporting News Baseball Guide* (1991), the probabilities of an out for these two pitchers are .739 and .646, respectively (computed from $1 -$ (hits + walks + hit batsmen)/batters faced). The parameter λ is more difficult to interpret, and corresponds roughly to the probability of facing exactly three batters in an inning, while at least one batter arrived safely at a base before being put out. To help in this interpretation a small simulation study was conducted based on 20 randomly selected games from the 1986 National League season, using Project Scoresheet account-form box scores (1986). These box scores provide a hard copy extended box score of each game in the season, allowing the user to reconstruct play-by-play information for each game, albeit in a less detailed fashion than the Project Scoresheet data tapes. During these 20 games there were 160 (41%) out of 388 innings (half-innings for visiting and home teams) in which exactly three batters were faced; in 22 (14%) of the 160 at least one batter either was credited with a hit or a walk or reached base safely (e.g., by a fielding error). In 12 of these 22 innings (55%) a double play subsequently occurred; in 7 of these 22 innings (32%) a runner was subsequently caught stealing, and in 3 (14%) of 22 innings a runner was put out on the base paths during the same at-bat (e.g., thrown out attempting to stretch a single to a double). Thus there is a substantial number of innings where three batters are faced and yet at least one batter reaches base safely (14%), thus obviating the need for the extra λ term in the specification of Equation (3). This term λ represents the ability of a pitcher (and catcher) to hold a runner close to first base, as well as the fielding ability of the team and the propensity of a pitcher to achieve double plays.

The parameters α and β relate to the probability that a batter will score a run given that he is a nonout. In particular, $\exp(\alpha + \beta)$ is the odds in favor of an individual batter scoring a run when exactly four batters have been faced in an inning; $\exp(\beta)$ is the odds ratio relating the probability of an individual batter scoring a run when $N + 1$ batters have been faced compared with the comparable probability when N batters have been faced. In general, high values of α and β indicate either poor pitcher performance

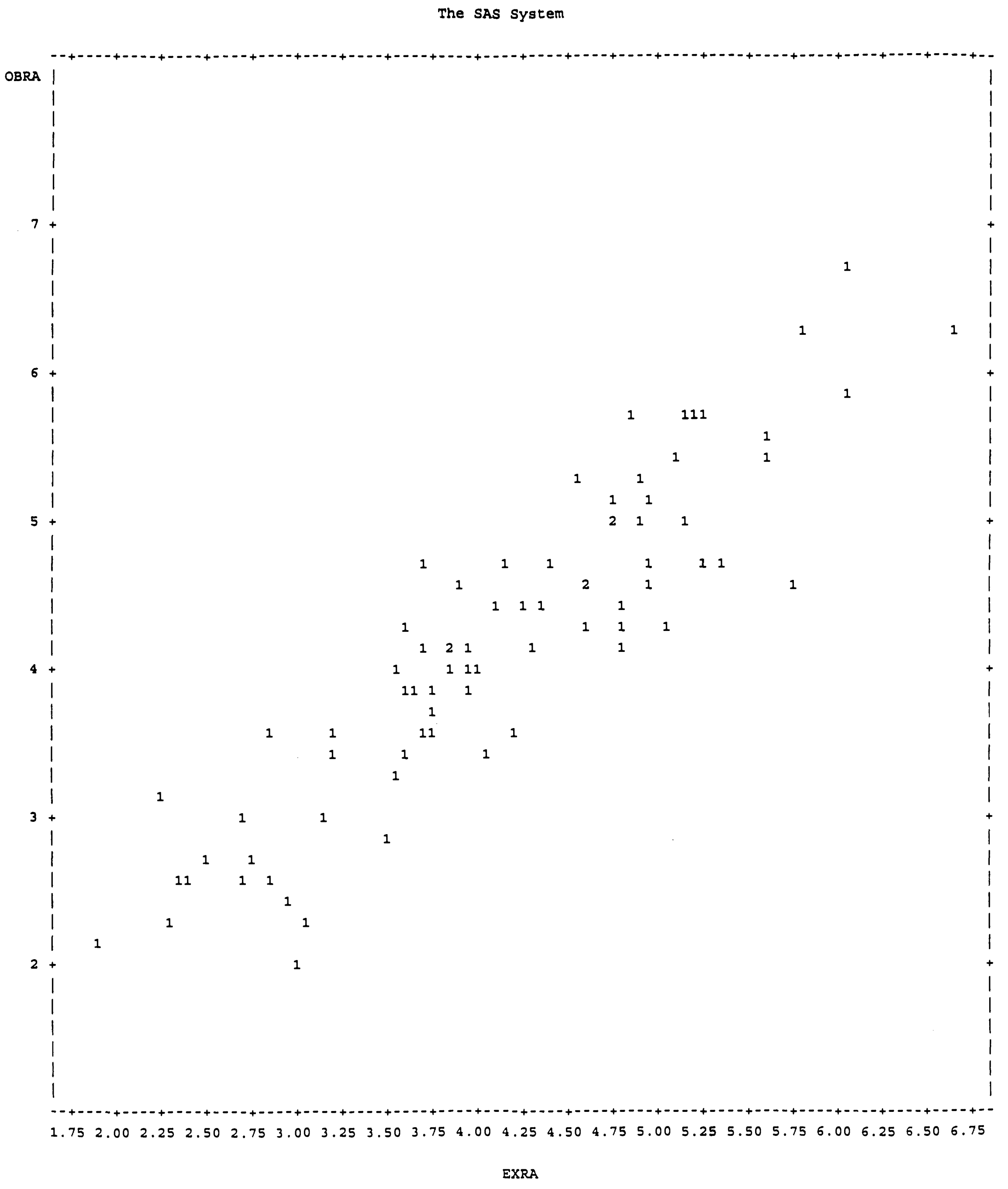

Figure 1. Scatterplot of OBRA Versus EXRA (77 Pitchers).

when men are already on base (i.e., poor "clutch" pitching) and/or a propensity to allow more home runs. However, it is difficult to identify marginal effects of these parameters because they are negatively correlated. A further complication is that the relative importance of α and β for an individual pitcher depends on the distribution of the number of batters faced. Specifically, if p, the probability of an out, is high, then α is likely to be of primary importance because the distribution of the number of batters faced has most of its probability mass at low values of N; conversely, if p is low, then both α and β have an important influence on the distribution of runs scored because the distribution of the number of batters faced has greater spread with a nontrivial probability mass at larger values of N.

An assumption made in fitting the models in this paper is that the sample of innings over all games for an individual pitcher consists of i.i.d. observations. Clearly, different batters are faced both in different games and even within different innings of the same game; thus the sample cannot consist of identically distributed random variables. As an example of the possible heterogeneity we computed the mean number of runs scored by inning over all pitchers (first inning, mean = .503 runs per inning, 1,905 innings; second inning, mean = .457 runs per inning, 1,888 innings; third inning, mean = .411 runs per inning, N = 1,846 innings). Thus there is some heterogeneity, even within different early innings of the same game. (By putting good batters high in the batting order, more runs are scored be-

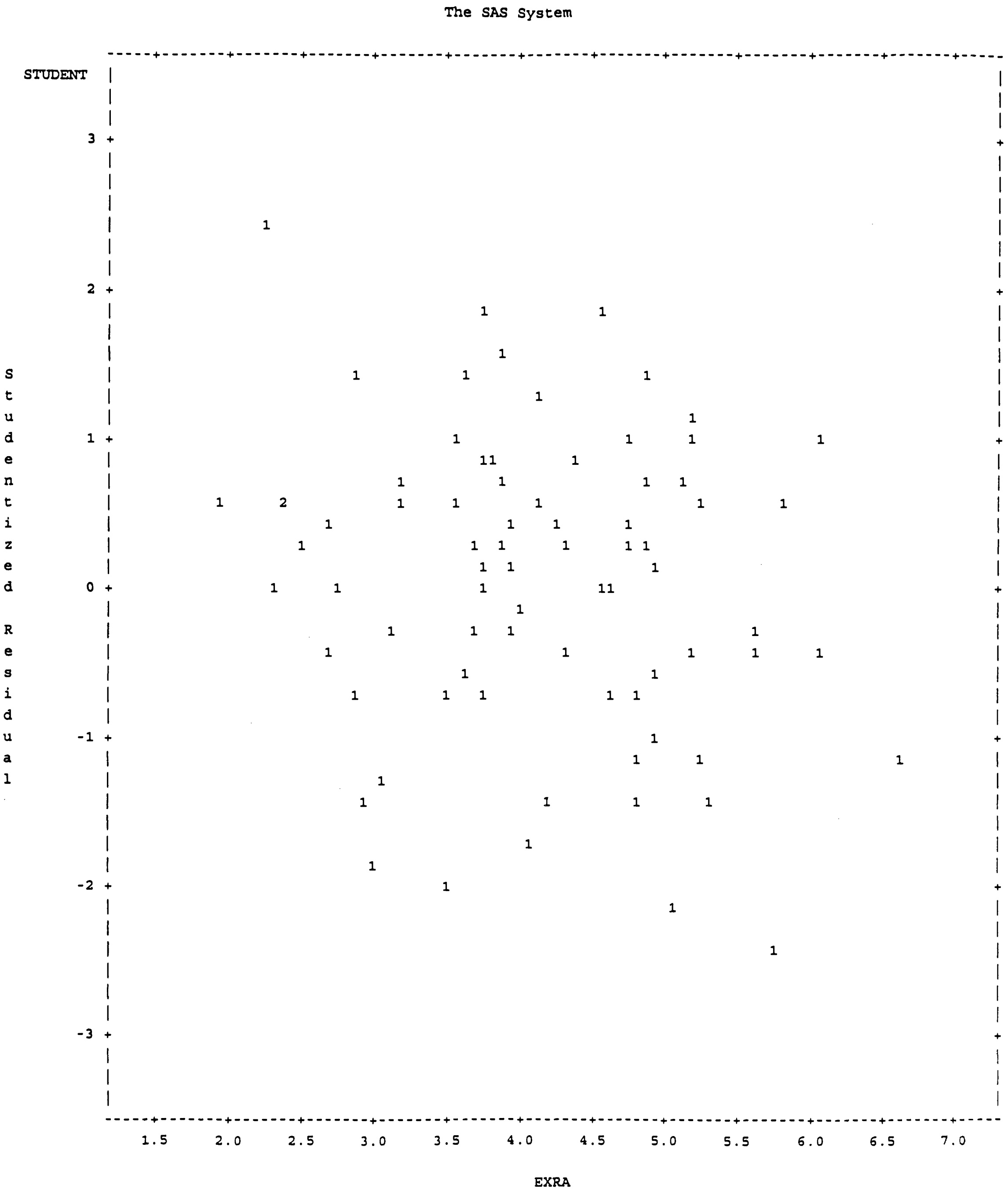

Figure 2. Plot of Studentized Residuals Versus EXRA (77 Pitchers).

cause they come to bat more often during a game.) Thus the distributions of runs scored and of batters faced are strictly applicable to an average batter. Further refinement of the modeling is necessary to accommodate individual differences in batting strengths and the effects of the order of the lineup.

Similarly, the assumption of independence presumes that pitcher performance is comparable in different games, that is, there is no correlation between the number of runs allowed and batters faced in different innings of the same game. The rationale for removing a pitcher in the middle of a game is often that the pitcher is having an "off day," which would imply a belief that pitcher performance in different games is substantially different. Nevertheless, it remains an open question whether for an individual pitcher, between game variation is substantially greater than variation between different innings of the same game. It is also unclear what the effect of violation of the i.i.d. assumption may have on parameter estimation. For example, in problems with correlated data, use of ordinary least squares instead of the preferable generalized least squares often results in unbiased estimation, but with standard errors that are biased downward (Liang and Zeger 1986).

In this article we have proposed a model for the distribution of the number of runs scored in innings of major league baseball games. The model is based on a convolution of the distribution of the number of batters faced in an inning and the conditional distribution of the number of runs scored given the number of batters faced, and it provided a numerically good fit to the observed distribution of runs scored, although not a large p value for the fit. The model is based on aggregate data for individual pitchers where the inning is the unit of analysis. An unsolved issue is how to simultaneously model both batter and pitcher performance in the same model. This may be possible using the Project Scoresheet database.

APPENDIX A: DERIVATION OF THE MLE FOR p, λ IN EQUATION (8)

Let I_N = number of innings for a particular pitcher, where the number of batters faced is N.

From Equation (3) the log likelihood is given by

$$\begin{aligned}\log L_1 &= \sum_{N=3}^{\infty} I_N \log[f_1(N)] \\ &= I_3 \log(p^3+\lambda) + I_4 \log[3p^3(1-p)-\lambda] \\ &\quad + \sum_{N=5}^{\infty} I_N \log\left\{\binom{N-1}{2}[p^3(1-p)^{N-3}]\right\}. \qquad \text{(A.1)}\end{aligned}$$

Upon differentiating (A.1) with respect to λ and setting $\partial \log L_1/\partial\lambda$ equal to 0 and solving for λ, we obtain the estimating equation when we replace λ by $\hat{\lambda}$ and p by $\hat{p}$ in

$$\lambda = \frac{3p^3(1-p)I_3 - p^3 I_4}{I_3+I_4}. \qquad \text{(A.2)}$$

By substituting (A.2) into (A.1) and simplifying we obtain

$$\begin{aligned}\log L_1 &= c + (I_3+I_4)\log[p^3+3p^3(1-p)] \\ &\quad + 3\sum_{N=5}^{\infty} I_N \log p + \sum_{N=5}^{\infty}(N-3)I_N \log(1-p), \qquad \text{(A.3)}\end{aligned}$$

where

$$c = I_3 \log\left[\frac{I_3}{I_3+I_4}\right] + I_4 \log\left[\frac{I_4}{I_3+I_4}\right].$$

We differentiate (A.3) with respect to p and obtain

$$\begin{aligned}\frac{\partial \log L_1}{\partial p} &= \frac{3\sum_{N=3}^{\infty} I_N}{p} - \frac{3(I_3+I_4)}{1+3(1-p)} \\ &\quad - \frac{\sum_{N=5}^{\infty}(N-3)I_N}{1-p}. \qquad \text{(A.4)}\end{aligned}$$

After setting the derivative (A.4) equal to 0 and solving the resulting quadratic equation we estimate $1-p$ as

$$\Delta = \frac{-c_1+c_2+2c_3 + \sqrt{(c_1-c_2-2c_3)^2+4c_3(3c_1+c_2+3c_3)}}{2(3c_1+c_2+3c_3)} \qquad \text{(A.5)}$$

and

$$c_1 = 3\sum_{N=3}^{\infty} I_N, \qquad c_2 = 3(I_3+I_4),$$

$$c_3 = \sum_{N=5}^{\infty}(N-3)I_N.$$

Equation (8) then follows from Equations (A.2) and (A.5).

Based on Equation (A.5), it can be shown that

$$\begin{aligned}\Pr(N=3) &= [\hat{p}^3 + 3\hat{p}^3(1-\hat{p})]\,\frac{I_3}{I_3+I_4} \\ \Pr(N=4) &= [\hat{p}^3 + 3\hat{p}^3(1-\hat{p})]\,\frac{I_4}{I_3+I_4}. \qquad \text{(A.6)}\end{aligned}$$

Thus the estimated probabilities $\Pr(N=3)$ and $\Pr(N=4)$ must always be between 0 and 1, inclusive.

APPENDIX B: PARAMETER ESTIMATES FOR INDIVIDUAL PITCHERS

Number of innings	p	λ	α	β	*OBRA*	*EXRA*	*Name*
98	.68587	.05118	−1.28852	.32030	4.77551	4.92998	1. Abbott, Jim
91	.71374	.08697	−2.34762	.68890	4.54945	3.89761	2. Anderson, Allan
72	.73456	.01440	−3.41019	.93368	2.75000	2.74042	3. Appier, Kevin
51	.75996	−.01302	−2.36959	.55598	2.29412	2.30646	4. Ballard, Jeff
92	.69304	.04448	−2.60603	.64897	4.20652	3.84190	5. Black, Bud
69	.70789	.06126	−2.16652	.62084	4.43478	4.10548	6. Blyleven, Bert
101	.71552	.04742	−1.71927	.47726	3.47525	4.04714	7. Boddicker, Mike
47	.74802	−.11232	−3.40767	.93671	2.68085	2.52228	8. Bolton, Tom
60	.71944	.01951	−2.40966	.69175	4.20000	3.72227	9. Bosio, Chris
76	.72260	.00675	−3.50878	1.08366	4.02632	3.53661	10. Brown, Kevin
87	.70806	−.02788	−2.83976	.75351	3.41379	3.60563	11. Candiotti, Tom
77	.72275	.00814	−2.43999	.97664	4.55844	4.93735	12. Cary, Charles
67	.69400	−.00589	−2.16085	.78004	5.50746	5.62200	13. Cerutti, John
92	.75925	−.02753	−3.69151	.94498	2.15217	1.91181	14. Clemens, Roger
60	.69633	−.03753	−2.57468	.61664	4.65000	3.72206	15. Davis, Storm
33	.71114	.09806	−1.01084	.47568	5.45455	5.61159	16. DuBois, Brian
51	.72025	.06542	−2.71100	.60131	3.00000	2.71064	17. Erickson, Scott
51	.73410	.02726	−1.19251	.16174	3.00000	3.15011	18. Farr, Steve
39	.68120	−.03500	−3.39832	1.00823	4.38462	4.78807	19. Fernandez, Alex
96	.73990	−.00792	−1.83194	.52768	3.28125	3.52952	20. Finley, Chuck

APPENDIX B: *(continued)*

Number of innings	p	λ	α	β	OBRA	EXRA	Name
95	.68387	.03001	−2.40898	.50098	3.78947	3.65997	21. Gordon, Tom
45	.69262	−.06799	−2.01359	.42343	4.20000	3.92998	22. Gubicza, Mark
62	.73179	−.01361	−2.76729	.89428	3.62903	3.69825	23. Guthrie, Mark
99	.75219	.05140	−2.32591	.68836	2.45455	2.94159	24. Hanson, Eric
91	.68191	.05914	−2.29279	.55624	4.45055	4.24053	25. Harnisch, Pete
89	.70193	.00430	−2.34829	.67505	4.14607	4.29962	26. Harris, Greg
73	.70423	.01263	−3.74826	1.40952	5.67123	5.16572	27. Hawkins, Andy
99	.73790	.05883	−2.37215	.75916	2.81818	3.52327	28. Hibbard, Greg
79	.70552	.02492	−3.55955	.91694	3.53165	3.20638	29. Higuera, Ted
84	.71298	.03001	−1.54727	.52549	4.28571	4.80955	30. Holman, Brian
96	.66016	.01357	−2.28041	.51787	5.15625	4.73971	31. Hough, Charles
34	.68352	.06604	−2.49764	.93498	5.82353	6.05893	32. Jeffcoat, Mike
87	.72207	.01803	−1.82109	.48631	3.51724	3.76364	33. Johnson, Dave
98	.69122	.04520	−1.70390	.33425	4.04082	3.92899	34. Johnson, Randy
70	.72642	.11520	−2.61680	.94454	3.60000	4.18151	35. Kiecker, Dana
75	.74550	−.04235	−2.92581	.71508	2.64000	2.39892	36. King, Eric
78	.69800	−.08589	−1.69317	.42909	4.61538	4.58698	37. Knudson, Mark
49	.67485	.10464	−.81933	.15709	4.95918	5.17013	38. Krueger, Bill
99	.72483	−.00256	−1.79751	.49931	4.00000	3.84426	39. Langston, Mark
81	.69870	−.00286	−1.69629	.57849	4.77778	5.32663	40. LaPoint, Dave
92	.70991	.00816	−2.64397	.80611	4.69565	4.15518	41. Leary, Tim
87	.73993	.02050	−2.18874	.53034	2.58621	2.85403	42. McCaskill, Kirk
45	.77445	−.11049	−3.05928	1.12266	2.00000	2.97937	43. McDonald, Ben
98	.66910	.06945	−1.70248	.39558	5.05102	4.89583	44. McDowell, Jack
32	.70110	.15339	−2.28536	.63568	3.93750	4.01668	45. McGaffigan, Andy
69	.68080	.01591	−2.51395	.75434	5.47826	5.10751	46. Milacki, Bob
50	.66825	.03237	−2.56539	.80145	4.50000	5.74760	47. Mitchell, John
98	.68433	.04431	−2.41520	.63255	4.77551	4.39904	48. Moore, Mike
108	.71957	.01159	−2.29645	.67464	4.16667	3.86193	49. Morris, Jack
30	.68026	−.02265	−1.49537	.34976	5.10000	4.94670	50. Moyer, Jamie
65	.66667	.08809	−2.14968	.49765	4.29231	4.59560	51. Navarro, Jaime
101	.69760	.05090	−1.71087	.57723	5.70297	5.20346	52. Perez, Melido
33	.71376	−.04178	−1.79162	.60001	4.09091	4.80974	53. Peterson, Adam
68	.67493	.08183	−1.90787	.45437	4.63235	4.61253	54. Petry, Dan
79	.66072	−.04270	−2.12110	.71012	6.26582	6.64875	55. Robinson, Jeff M.
66	.67028	.05503	−1.77562	.46971	4.77273	5.23492	56. Robinson, Ron
88	.71096	.04762	−1.52815	.52908	5.21591	4.89647	57. Ryan, Nolan
60	.72478	.05177	−2.62463	.62297	3.60000	2.84721	58. Saberhagen, Bret
102	.70332	.05061	−3.24693	1.17984	4.32353	5.03967	59. Sanderson, Scott
36	.75171	.04958	−1.63948	.43035	2.25000	3.05511	60. Searcy, Steve
69	.71585	.01726	−2.47009	.66725	3.91304	3.58907	61. Smith, Mike
108	.74743	−.00471	−2.52573	.56983	2.58333	2.36441	62. Stewart, Dave
96	.74516	−.02438	−3.38978	.84392	3.18750	2.22748	63. Stieb, Dave
99	.71405	.02965	−1.17626	.29989	4.45455	4.32703	64. Stottlemyre, Todd
100	.71241	.09435	−2.30254	.67282	3.87000	3.92926	65. Swindell, Greg
85	.70273	.11368	−1.05609	.30887	4.97647	4.74209	66. Tanana, Frank
82	.72024	.10496	−2.06705	.59818	3.84146	3.74941	67. Tapani, Kevin
36	.63371	.07193	−2.81077	.57571	5.00000	4.73722	68. Terrell, Walt
30	.70994	−.00729	−.33450	.16245	6.30000	5.80352	69. Tibbs, Jay
39	.64350	.13987	−3.13538	.90878	6.69231	6.06092	70. Valdez, Sergio
33	.64165	.00991	−2.15175	.46576	5.72727	5.26453	71. Walker, Mike
105	.73418	−.00856	−1.65253	.46618	3.77143	3.73977	72. Welch, Bob
74	.76025	.04133	−2.32942	.67436	2.55405	2.70691	73. Wells, David
79	.69771	.02016	−1.46304	.36848	5.35443	4.54668	74. West, David
96	.71974	−.05048	−2.84445	.73455	3.46875	3.21181	75. Witt, Bobby
46	.70202	−.02827	−2.17819	.69802	5.67391	4.85213	76. Witt, Mike
62	.71934	.00868	−2.21099	.59297	4.35484	3.62450	77. Young, Curt
5,639							

[Received January 1995. Revised July 1995.]

REFERENCES

Liang, K.-Y., and Zeger, S. L. (1986), "Longitudinal Data Analysis using Generalized Linear Models," *Biometrika*, 73, 13–22.

Project Scoresheet (1986), *1986 National League Account Form Box Scores*, Philadelphia: The Baseball Workshop.

——— (1990), *1990 American League Play-by-Play Data*, Philadelphia: The Baseball Workshop (619 Wadsworth Ave., Philadelphia, PA 19119).

The Sporting News (1991), *The Sporting News Baseball Guide—1991 Edition*, St. Louis: The Sporting News.