> **"Of the ninety-five best-of-seven World Series in baseball history through 2003, more series have gone the full seven games than any other length. Should this raise suspicions about the integrity of Major League Baseball?"**

Sporting Chances

Mark Schilling
California State University, Northridge

Though we often regard sporting events as tests of skill and valor, the element of chance plays a major role as well. Thus there are countless opportunities to apply probability and statistics to sports to test conjectures, analyze strategies, etc. Here are examples from two very popular sports:

Scoring in Soccer

Soccer is the world's favorite sport. Roughly two out of every three people on Earth are estimated to have watched at least part of the last World Cup tournament, held in Japan and Korea in 2002. Yet soccer has been criticized at times for lack of scoring. The paucity of goal scoring leads to frequent ties. Many people, though not necessarily soccer purists, consider a 0-0 game to be especially unappealing.

Beyond considerations of spectator enjoyment, in certain situations a winner must be chosen. Thus ties need to be broken. Sometimes an overtime period is played, very often resulting in no additional goals. If the contest is still tied after the overtime, the outcome may be decided by penalty kicks. In both the 1994 World Cup (men's) and the 1999 Women's World Cup, the championship game of the tournament was determined by penalty kicks. Though exciting, the penalty kick process—in which players from each team take turns taking shots at the opposite team's goalie—has been heavily criticized as an unsatisfactory way to choose a victor. (For one thing, teamwork is not involved.)

Attempts to tweak the rules to increase scoring in order to make ties and penalty kicks less frequent have so far been unsuccessful. The rule makers do not want to change the nature of a game that contains many subtle aspects.

Where the science of probability can help is in predicting the effects of scoring rates on the incidence of ties. We can do this by applying the following simple *Poisson process* model: assume that goal scoring occurs randomly in time, with the time of each goal being independent of the times of all other goals. Further assume that the *rate* of scoring does not change with time—i.e., that the probability of scoring a goal in the time interval $(t, t + \Delta t)$ does not depend on t. Then it can be shown that the probability that a team scores k goals within a period of time T is given by the Poisson distribution

$$P_T(k) = \frac{(\lambda T)^k e^{-\lambda T}}{k!}, \quad k = 0, 1, 2, \ldots$$

where λ is the average number of goals scored during a one-unit period of time.

A regulation soccer game lasts ninety minutes, plus perhaps five minutes on average of "injury time." With the Poisson model measuring T in minutes, the chance of a 0-0 tie is thus $P_{95}^2(0)$, the square occurring because both teams must fail to score. For games allowing a thirty-minute sudden death overtime, as in World Cup elimination rounds, the probability of a 0-0 game is approximately $P_{125}^2(0)$.

The model also estimates the likelihood of any kind of tie in a regulation game as $\sum_{k=0}^{\infty} P_{95}^2(k)$. (Do you see why?) The chance of a tie in a game with overtime is estimated by $\sum_{k=0}^{\infty} P_{95}^2(k) \times P_{30}^2(0)$, since the game must be tied at the end of regulation and then have no goals scored in the overtime.

Using the data from the 64 games of the 2002 World Cup, we find that 161 goals were scored in approximately 6200 minutes of playing time, or 0.0130 goals per team per minute. Taking 0.0130 as the value for λ, Table 1 displays the frequency of 0-0 scores, and of ties in general, from the formulas determined above.

	Prob(0-0 game)	Prob(tie)
Regulation game:	8.5%	27.2%
Playoff game with 30 min. sudden death overtime:	3.9%	12.5%

Table 1. The frequencies of 0-0 scores and ties using 2002 World Cup scoring rates.

Thus, at the rate of scoring seen in the 2002 World Cup, more than one out of four regulation games can be expected to end in a tie, and one out of eight playoff games will need to be decided by penalty kicks.

Applying the above percentages to the 2002 World Cup tournament, which consisted of forty-eight regulation games followed by sixteen playoff games, about four or five 0-0 games should have occurred. There were in fact three such games. The model also predicts about thirteen ties among the regulation games (27.2% × 48), and two ties in the elimination round (12.5% × 16), an almost perfect match to the fourteen ties in regulation games and two in playoff games that actually occurred.

Now let's see how much the frequencies of 0-0 games and ties are reduced if the rate of goal scoring is increased by 25%. Table 2 recomputes the above frequencies with $\lambda = 0.0130 \times 1.25$ goals per team per minute.

	Prob(0-0 game)	Prob(tie)
Regulation game:	4.6%	23.9%
Playoff game with 30 min. sudden death overtime:	1.7%	9.0%

Table 2. The frequencies of 0-0 scores and ties predicted assuming a 25% increase in the number of goals per minute.

The incidence of 0-0 games is cut roughly in half, while the decrease in the frequency of ties in general is much less dramatic. It seems that ties (and the likelihood of penalty kicks) are not easily avoided merely by a modest increase in scoring. You may wish to compute the effects on the above probabilities of doubling the goal scoring rate, a drastic change from soccer as it is currently played.

The Length of the World Series

Three of America's four major professional sports (baseball, basketball, and hockey) use a "best-of-seven" format for most rounds of their leagues' playoff tournaments. This means that the team that wins the majority of the seven games is declared the winner of the series. In practice, the series ends as soon as one team has won four games, since there is no need to play further games. Thus a series may last four, five, six, or seven games.

Occasionally when a series goes the full seven games, accusations are made that the series was extended on purpose, due to the wishes of the league and/or the television network covering the series, for the financial gain that results from having extra games. If there has indeed been a tendency for such conspiracies, it might not be provable in a specific instance but should be detectable in playoff data recorded over the years.

The final championship playoffs of Major League Baseball is the World Series. Of the ninety-five best-of-seven World Series in baseball history through 2003, more series have gone the full seven games than any other length. Should this raise suspicions about the integrity of Major League Baseball?

Suppose we consider the following simple model for the World Series: Assume that each game is like a coin flip—that is, the games are independent of each other and there is an equal chance for each team to win. The likelihood that a given World Series goes seven games is merely the chance that each team wins three of the first six games, a simple binomial probability. (The fact that some series last less than six games is irrelevant—can you see why?)

There are $\binom{6}{3} = 20$ ways to specify who wins each of the six games, each with probability $(1/2)^6$, thus the chance of a seven game series is $20 \times (1/2)^6 = 5/16$. We would therefore expect about $5/16 \times 95 = 29.7$ seven game series to have occurred, a bit less than the actual number, 35.

To properly judge the significance of this excess number of seven game series, we can do some additional analyses. By arguments similar to the one above, it is not hard to show (try!) that the full probability distribution of World Series length under our model is: Prob(4 games) = 1/8, Prob(5 games) = 1/4, Prob(6 games) = Prob(7 games) = 5/16. This is an example of a *negative binomial* distribution.

Table 3 compares the result of applying this to ninety-five series and the actual World Series records.

	4 games	**5 games**	**6 games**	**7 games**	**Average**
Actual	17	21	22	35	5.78
Model	11.9	23.8	29.7	29.7	5.81

Table 3. The predicted distribution for the lengths of ninety-five World Series compared to actual records for the lengths of the series.

We can see that there have been both more long (seven game) series *and* more *sweeps* (four game series) than the model predicts (and fewer six game series). As a result, the agreement between the actual *average* series length and the expected average series length for our model is excellent, giving some evidence that conspiracy theories about extending series on purpose are probably without merit.

What accounts for the discrepancies in the table above? One possibility is that the model we have used does not capture important elements affecting series length, such as differences in team ability, home field advantage, and dependency between game outcomes.

Intuitively the first of these factors, team ability, should result in more short series. If p is the chance in any game that

Continued on page 27

Continued from page 10

Team *A* beats Team *B*, then it is easy to see that Prob(4 games) $= p^4 + (1-p)^4 \geq 1/8$, achieving the minimum when $p = 0.5$ (representing teams with equal ability). Of course p varies from year to year, in fact from game to game, but it turns out that p has to stay quite far from 0.5 most of the time in order to greatly affect the frequency of sweeps. (As an exercise, find the two values of p that match $p^4 + (1-p)^4$ to the observed proportion of World Series sweeps.)

The home field advantage could conceivably account for frequent seven game series. This is because the World Series (as with the League Championship Series, and basketball and hockey playoffs) is structured so that, of the first six potential games, three take place at each team's home field. Now consider the case in which the home team nearly always wins. Clearly most series would last seven games. The actual home field advantage in baseball, though (around 57% of games are won by the home team), is not large enough to significantly affect the probability of a series going the maximum seven games.

The possibility remains that game outcomes are not independent of each other. For example, a team might gain confidence from winning one or two games and be more likely to prevail in subsequent contests. Extensive statistical studies of sports have not found evidence to refute the independence assumption, however.

The simplest explanation for the World Series data not matching our basic model better is simple chance variation. Data almost never fit a model perfectly. In Table 3, let the observed frequencies (the "Actual" row) and the expected frequencies (the "Model" row) be denoted by O_i and E_i, $i = 1, \ldots, 4$, respectively. The *chi-square goodness of fit test* computes

$$\chi^2 = \sum_{i=1}^{4} \frac{(O_i - E_i)^2}{E_i}$$

as a measure of the discrepancy between the data and the model. The resulting value is referenced to the chi-square probability distribution with $k-1$ degrees of freedom, where k is the number of categories of data—here, $k = 4$. It turns out that an overall discrepancy as large as the χ^2 value 5.46 found here is not particularly remarkable, having about a one out of seven chance of occurrence if the model is correct.

Viewed in this light, the World Series results above cannot be considered highly unusual. Thus there is no statistical reason to suspect Major League Baseball and/or the networks of influencing game outcomes to prolong the World Series. ■

Suggested Reading

H. Stern, Best-of-seven playoff series (A Statistician Reads the Sports Pages), *Chance Magazine*, 11 No. 2, Spring 1998.