# The Importance of a Game

MARK F. SCHILLING
California State University
Northridge, CA 91330

Organized play for many of the major sports that are popular around the world (baseball, basketball, hockey, tennis, and volleyball, for example) involves a *contest* between two opponents in which the victor is the team or individual that is the first to win a prescribed number of subcontests. Often during these contests one will hear in conversation or on a broadcast a statement such as "This game is the most important game of the series". Why are some games perceived as being more important than others, when each game counts the same toward the total number of victories required to win the contest?

The purpose of this article to analyze the meaning and validity of such statements from a probabilistic perspective and to provide a means of quantifying importance, in order that the comparative significance of games can be measured. In this paper, the word *game* will refer to the subcontests that comprise a contest, and one participant in the contest will be arbitrarily denoted "the contestant", or "C". The key observation is that the importance of a game depends on context. Two distinct versions of importance, developed below, are therefore appropriate. It will be assumed that the outcome of each game in a contest is independent of all other outcomes. Participants in such contests frequently dispute the reasonableness of this supposition, citing the effects of momentum and other factors, primarily psychological; however, evidence from statistical analysis of contests [2] suggests that actual experience is quite compatible with the independence assumption.

**Conditional and a priori importance** To fix notation, let $g$ be the number of a particular game in the sequence of games comprising a contest, and suppose that contestant C has won $k$ of the previous $g-1$ games.

*Definition* 1. The *conditional importance* $I(g|k)$ of game $g$ is the difference between (1) the conditional probability that C wins the contest if C *wins* the game and (2) the conditional probability that C wins the contest if C *loses* the game, given the status of the contest up to game $g$.

Note that $I(g|k)$ is always nonnegative and is the same for both contestants. Here is an example: Suppose two teams are playing a series in which the first team to win four games (commonly known as a "best of seven" series) emerges victorious. Assume that four games have been played so far, with each team having earned two wins. Then the conditional importance of the fifth game is

$$\begin{aligned} I(5|2) &= \mathrm{P}(\text{C wins series}|\text{C wins game 5}) - \mathrm{P}(\text{C wins series}|\text{C loses game 5}) \\ &= \mathrm{P}(\text{C wins series}|\text{C leads 3 games to 2}) \\ &\quad - \mathrm{P}(\text{C wins series}|\text{C trails 2 games to 3}) \\ &= \mathrm{P}(\text{C wins at least 1 of the 2 remaining games}) \\ &\quad - \mathrm{P}(\text{C wins both remaining games}). \end{aligned}$$

If each team is equally likely to win each game independently of the outcomes of any of the previous games, then the conditional importance of game 5 is

$$I(5|2) = 3/4 - 1/4 = 1/2.$$

That is, since the contestant's chances of winning the series will be 75% if game 5 is won and 25% if it is lost, the conditional importance of game 5 is 75% − 25% = 50%.

If the final game of a contest with a limited number of games (such as the seventh game of a "best of seven" series), is played, it clearly has the highest possible conditional importance, namely 1 − 0 = 100%. No other game can have a conditional importance as high as this, provided each team has a positive probability of winning each game.

Why then, are other games ever considered the "most important game"? The fifth game in a "best of seven" series—the example shown above—is often described this way, as is the seventh game in a set in tennis. And can it really be true that one game is any more or less important than any other?

The answers to these questions depend on the fact that a game will have no importance at all if it is not played. For example, the seventh game of a "best of seven" series is played only if the first six games are split evenly, with three wins for each contestant. Otherwise the conditional importance of game seven is zero—even if the game were played it could not change the winner of the contest. It can therefore be argued that importance should take into account the chance that the game will be necessary, and more generally, the possible states of the contest when that game is played (if at all), as well as the conditional importance of the game given these possible states.

*Definition* 2. The *a priori importance* of game $g$ is $I(g) = EI(g|k)$ = the *expected* conditional importance of the game.

Conditional importance and a priori importance will clearly coincide for game $g = 1$. We shall see, however, that in general their forms are quite different.

Assume henceforth that to win a contest it is necessary to win a fixed number $n > 1$ of games, with at most $2n - 1$ games necessary in all, although the same ideas can easily be extended to other types of contests. Note that in the example above, the conditional importance of game 5 is simply P(C wins *exactly* 1 of the remaining 2 possible games). In general,

$$\begin{aligned} I(g|k) &= \text{P(C wins at least } n-k-1 \text{ of the remaining possible games)} \\ &\quad - \text{P(C wins at least } n-k \text{ of the remaining possible games)} \\ &= \text{P(C wins exactly } n-k-1 \text{ of the } 2n-1-g \text{ possible games} \\ &\qquad \text{remaining after game } g). \qquad (1) \end{aligned}$$

Now let $P_g(k)$ be the probability that $C$ wins $k$ of the first $g - 1$ games. Then from Definition 2,

$$\begin{aligned} I(g) &= \sum_{k=0}^{g-1} I(g|k) P_g(k) \\ &= \sum_{k=0}^{g-1} \text{P(C wins } n-k-1 \text{ games after game } g) \times \\ &\qquad \text{P(C wins } k \text{ games before game } g) \\ &= \text{P(C wins } n-1 \text{ of the games other than game } g). \qquad (2) \end{aligned}$$

Note that the event in (2) is that $C$ wins exactly half of the $2n - 2$ games other than game $g$. This makes sense intuitively since, if the games other than game $g$ result in an equal number of victories for each participant, then game $g$ represents a "decisive game". With this interpretation, any contest will turn out to have either $n$

decisive games or *no* decisive games, depending on whether the contest does or does not extend to its last possible game. Furthermore, using (2) we can represent the a priori importance as

$$I(g) = \mathrm{P}(\mathrm{C \text{ wins the contest}} | \mathrm{C \text{ wins game} } g) - \mathrm{P}(\mathrm{C \text{ wins the contest}} | \mathrm{C \text{ loses game} } g),$$

which is the *unconditional* parallel to conditional importance.

Two simple models are appropriate for many contests of the "majority wins" type described above.

**Model 1.** The games follow a binomial model in which the win probability $0 < p < 1$ for contestant C is the same for each game.

**Model 2.** The contest consists of independent games that are of two types, say "type 1" and "type 2", each following a binomial model with win probabilities $0 < p_1$, $p_2 < 1$ respectively.

Model 2 reduces to Model 1 when $p_1 = p_2$. Model 1 is appropriate when the conditions are essentially the same for each game. Model 2 is representative of sporting contests in which the games are played at two sites, normally the home towns of the two contestants, with type 1 and type 2 being the games played at these two sites. This model would also be appropriate in tennis (volleyball) when a set represents a contest, since the win probability for each game (point) often depends heavily on who is serving.

First let us analyze the two importance quantities under Model 1. Note that although a contest will normally terminate as soon as the requisite number of games is won by one of the participants, all analyses can be carried out as if all $2n-1$ games are actually played, since the winner is the same in either case. Define

$$\binom{u}{v} = \begin{cases} \dfrac{u!}{v!(u-v)!} & \text{for } 0 \le v \le u; \\ 0 & \text{otherwise.} \end{cases}$$

Since at most $2n-1$ games can be played, equation (1) gives for the conditional importance

$$I(g|k) = \binom{2n-1-g}{n-k-1} p^{n-k-1} q^{n-g+k}, \tag{3}$$

where $q = 1-p$.

We can compute the a priori importance by making use of the following combinatorial result (see [1], p. 64), which will be referred to as the *hypergeometric identity* because of its direct relationship to the hypergeometric probability distribution:

$$\sum_{k=0}^{m} \binom{a}{k}\binom{b}{m-k} = \binom{a+b}{m}.$$

This identity can be seen to count the number of ways to choose $m$ items from a set of $a+b$ items of which $a$ are of one kind and $b$ are of another.

Now since $P_g(k) = \binom{g-1}{k} p^k q^{g-1-k}$, the a priori importance for Model 1 is

$$I(g) = \sum_{k=0}^{g-1} \binom{2n-1-g}{n-k-1}\binom{g-1}{k} p^{n-1} q^{n-1}$$

$$= \binom{2n-2}{n-1} p^{n-1} q^{n-1}, \tag{4}$$

upon using the hypergeometric identity. The a priori importance is *the same* for each game under Model 1. Note that its value is also the probability that the contest will extend to the last game.

Under Model 2, the conditional and a priori importances are obtained from convoluting the two distinct binomial distributions involved. Write $q_i = 1 - p_i$ and let $n_i$ represent the number of games of type $i$ remaining after game $g$, for $i = 1, 2$. Then the conditional importance of game $g$ under Model 2 is

$$I(g|k) = \sum_{j=0}^{n_1} \binom{n_1}{j} p_1^j q_1^{n_1-j} \binom{n_2}{n-k-1-j} p_2^{n-k-1-j} q_2^{n_2-n+k+1+j}. \tag{5}$$

The a priori importance is *not* the same for each game under Model 2, since $I(g)$ depends on the total number of possible games $t_i$ of each type, $i = 1, 2$, other than game $g$, as well as on the win probabilities. Decomposing over $j =$ the number of type 1 games other than game $g$ that are won by C yields for (2)

$$I(g) = \sum_{j=0}^{t_1} \binom{t_1}{j} \binom{t_2}{n-1-j} p_1^j q_1^{t_1-j} p_2^{n-1-j} q_2^{t_2-n+1+j}. \tag{6}$$

This result can also be computed directly from the definition of a priori importance by repeated application of the hypergeometric identity. Looking at (6), we see that since game $g$ is itself one of the two types of games, but is not counted by either $t_1$ or $t_2$, all games of type 1 must have equal a priori importance, as must all games of type 2; however these two values normally will not be equal to each other.

**Analysis of the importance functions** The degree of conditional importance attached to a game depends on the extent to which the ultimate outcome of the contest remains in doubt at that point. To gain insight into this phenomenon, let us examine $I(g|k)$ for Model 1 more closely. Let $N = 2n - 1 - g$, the number of possible games remaining after game $g$, and write $m = n - k - 1$ for the number of additional wins the contestant C must attain to win the contest if game $g$ is also won. The conditional importance (3) then takes the simple binomial form

$$I(g|k) = \binom{N}{m} p^m q^{N-m}.$$

Using standard properties of the binomial distribution, we find that for given $N$ and $p$, $I(g|k)$ is a unimodal function on $m = 0, 1, \ldots, n$ reaching its maximum value at $m = [(N+1)p]$, where $[x]$ represents the greatest integer less than or equal to $x$. Similarly, as a function of $p$ with $N$ and $m$ fixed, $I(g|k)$ is maximized at $p = m/N$ (by elementary calculus). From either perspective, conditional importance is greatest when the proportion of the possible remaining games that C must win is at or near to C's win probability for a single game. Comparable behavior occurs under Model 2 unless the values of $p_1$ and $p_2$ are very different from each other.

Natural questions to ask about a priori importance are:

(1) When is $I(g)$ small? large?
(2) When are games of one type (e.g., "home" games) more or less important than games of the other type ("away" games)?

The first question is easier to answer than the second: $I(g)$ can be made arbitrarily close to 0—for example, if $p_1$ and $p_2$ tend to 1. $I(g)$ can also become arbitrarily close to 1 if, for instance, $t_1 = t_2$ and $p_1$ tends to 1 while $p_2$ tends to 0.

With regard to the second question, if $p_1 = p_2$ then all games are equally important. In general, it turns out that the a priori importance of type 1 games can be greater than, lower than, or equal to that of type 2 games, regardless of which type comprises the majority of games in the contest.

Consider those contests in which one more game is scheduled of one type, say type 1, than of the other type. This is the usual setup in sporting contests; for example, the World Series in Major League Baseball has four games in one stadium and three in another, if the Series extends to the maximal number of games. If game $g$ is of type 1, then $t_1 = t_2 = n - 1$, while if game $g$ is of type 2, then $t_1 = n$ and $t_2 = n - 2$. Denote the a priori importances for type 1 and type 2 games by $I_1$ and $I_2$, respectively. Although a simple analytical result appears unobtainable, numerical studies indicate an interesting relationship, which is summarized in FIGURES 1, 2, and 3 below. These graphs indicate the sign of $\Delta I = I_1 - I_2$ for "best of three" ($n = 2$), "best of five" ($n = 3$) and "best of seven" ($n = 4$) contests, respectively, as a function of $p_1$ and $p_2$. In each case, both positive and negative values of $\Delta I$ can occur.

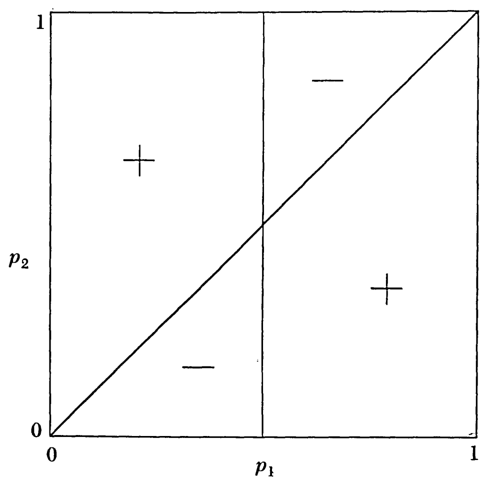


**FIGURE 1**
Sign of $\Delta I$ for "best of three" contests.

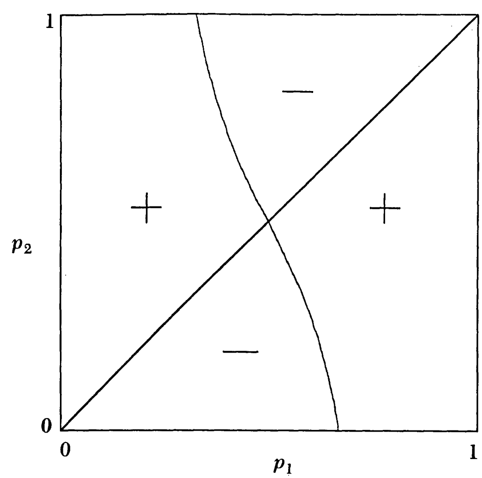


**FIGURE 2**
Sign of $\Delta I$ for "best of five" contests.

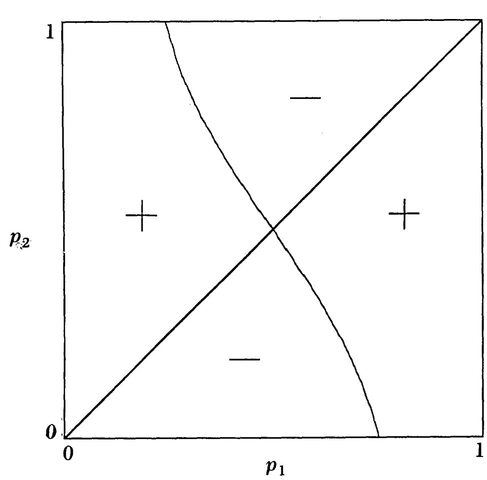


**FIGURE 3**
Sign of $\Delta I$ for "best of seven" contests.

Specifically, suppose that type 1 games are the home games for contestant C and type 2 games are the games played away from home. Invariably this means that $p_1$ will be greater than $p_2$, a fact generally known as the "home field (or court) advantage". In this setting, the home games will be *more* important than the away games if $p_1 > p_2 > .5$, that is, if C is favored to win each game in the contest, while the home games will be *less* important than the away games if C is the underdog at both sites. If each participant is favored at its own site, however, the relative a priori importance depends further on $p_1$ and $p_2$ as well as on $n$, with the home games of C being the more important ones if $p_1$ is sufficiently large compared to $p_2$.

The following example indicates the reasonableness of the result: Suppose a "best of three" contest is to be played and that each team is heavily favored to win at its home site, with two games to be played at the home site of the contestant C and one at the opponent's home site. According to FIGURE 1, C's home games are the more important ones. This makes sense intuitively because C should win the contest regardless of the outcome of the away game. Thus that game is relatively unimportant, while C's home games are quite important because losing a home game would decrease C's chances of winning the contest from near 1 to near 0.

**Application: the 1988 NBA Championship Series** To illustrate the concepts of conditional and a priori importance in a real world setting, the values of these quantities are estimated for the 1988 National Basketball Association (NBA) Championship Series, which was played between the Los Angeles Lakers and the Detroit Pistons. The series was a "best of seven" affair, with four games held in Los Angeles and three held in Detroit. Los Angeles will therefore be the contestant C in this application, and games in L.A. will be the type 1 games while type 2 games will represent those in Detroit. It is necessary to estimate $p_1$ and $p_2$ in order to estimate the importance values. The following procedure leads to reasonable results:

1. First estimate the *odds* $o(L.A.)$ in favor of Los Angeles defeating Detroit in a single game at a neutral site from the season records of each team as follows: L.A. won 62 games and lost 20 in 1987–88, Detroit won 54 and lost 28, so we estimate $o(L.A.)$ by

$$\hat{o}(L.A.) = \frac{62}{20} \times \frac{28}{54} = 1.607.$$

(This method has been suggested by several authors; see [3] and [4] for details and references.)

2. Data from NBA games indicate that the home team wins approximately twice as often as the away team, so we modify $\hat{o}(L.A.)$ to obtain home odds $\hat{o}_1(L.A.) = 2 \times \hat{o}(L.A.) = 3.215$ and away odds $\hat{o}_2(L.A.) = .5 \times \hat{o}(L.A.) = 0.804$. This leads to the following estimates for $p_1$ and $p_2$:

$$\hat{p}_1 = \frac{\hat{o}_1(L.A.)}{\hat{o}_1(L.A.) + 1} = .763 \quad \text{and} \quad \hat{p}_2 = \frac{\hat{o}_2(L.A.)}{\hat{o}_2(L.A.) + 1} = .446.$$

Table 1 provides the schedule of games, the results, and the values of $I(g)$ and $I(g|k)$ for the series, computed for Model 2 from these estimates of $p_1$ and $p_2$. For comparison, conditional importance values are also given for Model 1 with $p = .5$, which would apply if each team were equally likely to win each game.

Observe in Table 1 that the games in Los Angeles had greater a priori importance than those in Detroit, in agreement with FIGURE 3 for the estimates of $p_2$ and $p_2$ given above. This can be understood intuitively by noting that although L.A. was favored to win the series, a loss of just one of their home games would have been decisive if each other game were won by the home team. The games in Detroit were less important because, as long as Los Angeles won at home, the games in Detroit could not change the outcome of the series. The oft-stated "importance of holding onto the home-court advantage" is therefore seen to be well justified in this case.

TABLE 1. A Priori and Conditional Importances, 1988 NBA Championship Series

| Game | Site | $I(g)$ | $I(g\|k)$ | $I(g\|k)$ (Model 1) | Winner |
|---|---|---|---|---|---|
| 1 | L.A. | .289 | .289 | .312 | Detroit |
| 2 | L.A. | .289 | .364 | .312 | L.A. |
| 3 | Detroit | .235 | .369 | .375 | L.A. |
| 4 | Detroit | .235 | .226 | .375 | Detroit |
| 5 | Detroit | .235 | .362 | .500 | Detroit |
| 6 | L.A. | .289 | .763 | .500 | L.A. |
| 7 | L.A. | .289 | 1.000 | 1.000 | L.A. |

Since this series went to the maximum number of games, conditional importance naturally increased dramatically toward the end. It is interesting that the fourth game turned out to be the game with the lowest conditional importance. The reason is that Los Angeles was in a very strong position at that time, ahead two games to one—they would have been heavily favored to win the series *regardless* of the outcome of game 4. Note also that the fifth game did not have critical importance (for similar reasons), despite the traditional dogma about such games mentioned earlier.

**Concluding remarks** There may be at least one practical application of the concepts of a priori and conditional importance. If the audiences for sporting contests tend to sense subjectively the approximate relative importances of the games, then advance calculations could provide a valuable means of predicting audience interest for television and radio networks, ticket sellers, and promoters. Rates for advertising time, for example, might be set in part according to the estimated a priori importance of a contest's games, or based on conditional importance (if deadline considerations permit) according to the progress of a series. Studies of past broadcast ratings and future surveys would be of interest in order to determine whether such connection exists.

The notions presented here have been illustrated only for contests of a specific kind, where the requirement is to win a predetermined number of games before the adversary does. However, the definitions and techniques can be adapted easily to any situation where the ultimate outcome depends probabilistically on the occurrence of a number of component events.

## REFERENCES


1. W. Feller, *An Introduction to Probability Theory and Its Applications*, Vol. 1, 3rd. edition, John Wiley & Sons, Inc., New York, 1968.
2. T. Gilovich, R. Vallone and A. Tversky, The hot hand in basketball: On the misperception of random sequences, *Cogn. Psych.* 17 (1985), 295–314.
3. T. Jech, The ranking of incomplete tournaments: a mathematician's guide to popular sports, *Amer. Math. Monthly* 90 (1983), 246–266.
4. M. Stob, A supplement to "A mathematician's guide to popular sports", *Amer. Math. Monthly* 91 (1984), 277–282.