THE JAMES FUNCTION

CHRISTOPHER N. B. HAMMOND, WARREN P. JOHNSON, AND STEVEN J. MILLER

Abstract. We investigate the properties of the James function, associated with Bill James's so-called "log5 method," which assigns a probability to the result of a game between two teams based on their respective winning percentages. We also introduce and study a class of functions, which we call *Jamesian*, that satisfy the same *a priori* conditions that were originally used to describe the James function.

1. Introduction

In his 1981 *Baseball Abstract* [8], Bill James posed the following problem: suppose two teams A and B have winning percentages a and b respectively, having played equally strong schedules in a game such as baseball where there are no ties. If A and B play each other, what is the probability $P(a, b)$ that A wins?

This question is perhaps more relevant to other sports, because in baseball the outcome is particularly sensitive to the pitching matchup. (In 1972, the Philadelphia Phillies won 29 of the 41 games started by Steve Carlton, and 30 of the 115 games started by their other pitchers.) The answer is quite interesting, even if its applicability is somewhat limited by the tacit assumption of uniformity.

For $0 < a < 1$ and $c > 0$, define $q_c(a)$ by

$$a = \frac{q_c(a)}{q_c(a) + c}. \tag{1.1}$$

James calls $q_{\frac{1}{2}}(a)$ the log5 of a, and does not consider any other values of c. He claims that

$$P(a, b) = \frac{q_{\frac{1}{2}}(a)}{q_{\frac{1}{2}}(a) + q_{\frac{1}{2}}(b)}. \tag{1.2}$$

This technique is sometimes called the log5 method of calculating $P(a, b)$, although we will avoid using this name as there is nothing obviously logarithmic about it. It is easy to see from (1.1) that

$$q_c(a) = \frac{ca}{1 - a}.$$

Substituting this expression into (1.2), we get

$$P(a, b) = \frac{a(1 - b)}{a(1 - b) + b(1 - a)},$$

not only for $c = \frac{1}{2}$ but for any positive c.

Date: January 13, 2014.

Key words and phrases. James function, Bradley–Terry model, sabermetrics.

The third-named author was partially supported by NSF grant DMS1265673.

The explicit form of $P(a,b)$ was first given by Dallas Adams [8], who also christened it the *James function.* It makes sense to extend the James function to values of a and b in the set $\{0,1\}$, except when $a = b = 0$ or $a = b = 1$. In these two cases, we would not have expected to be able to make predictions based on winning percentages alone. Moreover, both cases would be impossible if the two teams had previously competed against each other.

James's procedure can be interpreted as a twofold application of the general method known as the *Bradley–Terry model* (or sometimes the *Bradley–Terry–Luce model*). If A and B have worths $w(A)$ and $w(B)$ respectively, the probability that A is considered superior to B is

$$\pi(A,B) = \frac{w(A)}{w(A)+w(B)}.$$

Despite the attribution of this model to Bradley and Terry [1] and to Luce [10], the basic idea dates back to Zermelo [15]. The question, of course, is how to assign the "right" measure for the worth of A in a particular setting. In chess, for instance, it is common to express the worth of a player as $10^{R_A/400}$, where R_A denotes the player's Elo rating (see [5]). (The rating of chess players is the question in which Zermelo was originally interested. Good [6], who also considered this problem, seems to have been the first to call attention to Zermelo's paper.) Another example is James's so-called Pythagorean model (introduced in [7, p. 104] and discussed further in [12]) for estimating a team's seasonal winning percentage, based on the number R of runs it scores and the number S of runs it allows. In this case, the worth of the team is R^2 and the worth of its opposition is S^2.

In the construction of the James function, we can view the measure of a team's worth as being obtained from the Bradley–Terry model itself. We begin by assigning an arbitrary worth $c > 0$ (taken by James to be $\frac{1}{2}$) to a team with winning percentage $\frac{1}{2}$. Equation (1.1) can be construed as an application of the Bradley–Terry model, where the worth of a team is determined by the fact that its overall winning percentage is equal to its probability of defeating a team with winning percentage $\frac{1}{2}$. Equation (1.2) represents a second application of the Bradley–Terry model, where each team has an arbitrary winning percentage and the measure of its worth comes from the previous application of the model.

This area of study, which is usually called the theory of paired comparisons, has focused from the outset [15] on the question of inferring worth from an incomplete set of outcomes. (See [2] for a thorough treatment, as well as [4] and [14] for additional context.) James, on the other hand, takes the worths to be known and uses them to determine the probability of the outcomes. We will adopt a similar point of view, emphasizing a set of axiomatic principles rather than a specific model.

James's justification [8] for his method does not invoke the Bradley–Terry model, but rather the fact that the resulting function $P(a,b)$ satisfies six self-evident conditions:

(1) $P(a,a) = \frac{1}{2}$.
(2) $P(a,\frac{1}{2}) = a$.
(3) If $a > b$, then $P(a,b) > \frac{1}{2}$, and if $a < b$ then $P(a,b) < \frac{1}{2}$.
(4) If $b < \frac{1}{2}$, then $P(a,b) > a$, and if $b > \frac{1}{2}$ then $P(a,b) < a$.
(5) $0 \le P(a,b) \le 1$, and if $0 < a < 1$ then $P(a,0) = 1$ and $P(a,1) = 0$.

(6) $P(a, b) + P(b, a) = 1$.

To avoid contradicting (5), condition (4) should exclude the cases where $a = 0$ and $a = 1$. We will call this set, with this slight correction, the *proto-James conditions*. In addition to presenting some empirical evidence for (1.2), James makes the following conjecture.

Conjecture 1 (James, 1981)**.** *The James function $P(a, b)$ is the only function that satisfies all six of the proto-James conditions.*

Jech [9] independently proposed a similar, albeit shorter list of conditions. Although he did not consider Conjecture 1, he was able to prove a uniqueness theorem pertaining to a related class of functions.

While the proto-James conditions are certainly worthy of attention, we prefer to work with a slightly different set. The following conditions pertain to all points (a, b) with $0 \le a \le 1$ and $0 \le b \le 1$, except for $(0, 0)$ and $(1, 1)$:

(a) $P(a, \frac{1}{2}) = a$.
(b) $P(a, 0) = 1$ for $0 < a \le 1$.
(c) $P(b, a) = 1 - P(a, b)$.
(d) $P(1 - b, 1 - a) = P(a, b)$.
(e) $P(a, b)$ is a strictly increasing function of a, unless $b = 0$ or $b = 1$.

We shall refer to conditions (a) to (e) as the *James conditions*.

It is fairly obvious that the James conditions imply the proto-James conditions. Condition (a) is identical to condition (2). Condition (c) is condition (6), which implies (1) by taking $b = a$. Condition (d) simply says that the whole theory could be reformulated using losing percentages instead of winning percentages, with the roles of the two teams reversed. Condition (e) is stronger than (3) and (4), and in concert with (1) and (2) implies them both. Combined with (c) or (d), it also implies that $P(a, b)$ is a strictly decreasing function of b unless $a = 0$ or $a = 1$. Finally, (b) implies the second of the three parts of (5). Together with (c), it also implies that $P(0, b) = 0$ if $0 < b \le 1$. By taking $b = 0$ in (d) and replacing $1-a$ with b, condition (b) further implies that $P(1, b) = 1$ if $0 \le b < 1$, and this together with (c) gives $P(a, 1) = 0$ for $0 \le a < 1$, which is (a hair stronger than) the third part of (5). These facts, combined with (e), show that $0 < P(a, b) < 1$ when $0 < a < 1$ and $0 < b < 1$, which implies the first part of (5).

The purpose of this paper is to examine the mathematical theory underlying the James function and to demonstrate that Conjecture 1 is actually false. In fact, we will introduce and study a large class of functions that satisfy the James conditions (and hence the proto-James conditions).

2. Verification of the James Function

While the Bradley–Terry model is practically ubiquitous, it is not obvious from an axiomatic point of view. We will now present a self-contained proof that the James function represents the probability $P(a, b)$, under the assumption that the winning percentage of each team provides an exact measure of its current quality. The following argument was discovered by the third-named author several years ago [11], but has not previously appeared in a formal publication.

Theorem 2. *The probability $P(a,b)$ that a team with winning percentage a defeats a team with winning percentage b is given by the James function*

$$\frac{a(1-b)}{a(1-b)+b(1-a)}, \tag{2.1}$$

except when $a=b=0$ or $a=b=1$, in which case $P(a,b)$ is undefined.

Proof. Let teams A and B have winning percentages a and b respectively. Independently assign to each of A and B either a 0 or 1, where A draws 1 with probability a and B draws 1 with probability b. If one team draws 1 and the other 0, then the team with 1 wins the game. If both teams draw the same number, then repeat until they draw different numbers.

The probability that A draws 1 and B draws 0 on any given turn is clearly $a(1-b)$, while the opposite occurs with probability $b(1-a)$. The probability that A and B both draw 1 is ab, and the probability that they both draw 0 is $(1-a)(1-b)$. Hence

$$ab+(1-a)(1-b)+a(1-b)+b(1-a)=1. \tag{2.2}$$

It follows that $0 \leq ab+(1-a)(1-b) \leq 1$ and $0 \leq a(1-b)+b(1-a) \leq 1$ whenever $0 \leq a \leq 1$ and $0 \leq b \leq 1$.

We can conclude the argument in either of two ways. Since the probability that A and B draw the same number is $ab+(1-a)(1-b)$, in which case they draw again, $P(a,b)$ must satisfy the functional equation

$$P(a,b)=a(1-b)+[ab+(1-a)(1-b)]\,P(a,b).$$

The only case in which we cannot solve for $P(a,b)$ is when $ab+(1-a)(1-b)=1$. By (2.2), this situation only occurs when $a(1-b)+b(1-a)=0$, which implies that either $a=b=0$ or $a=b=1$. Otherwise we obtain the stated formula for $P(a,b)$.

Alternatively, we may observe that the probability that A wins on the nth trial is

$$a(1-b)\,[ab+(1-a)(1-b)]^{n-1},$$

and so the probability that A wins in at most n trials is

$$a(1-b)\sum_{k=1}^{n}[ab+(1-a)(1-b)]^{k-1}.$$

As n tends to ∞, this expression yields a convergent geometric series unless $ab+(1-a)(1-b)=1$. Using (2.2), we again obtain the James function. □

It should be clear from (2.1) that the James function satisfies James conditions (a) to (d). We will verify condition (e) momentarily.

3. Properties of the James function

In this section, we will consider several important properties of the James function. We begin by computing the partial derivatives of $P(a,b)$, which will lead to an observation originally due to Dallas Adams. Note that

$$\frac{\partial P}{\partial a}=\frac{b(1-b)}{[a(1-b)+b(1-a)]^2}, \tag{3.1}$$

which shows that the James function satisfies condition (e), and also

$$\frac{\partial P}{\partial b} = \frac{-a(1-a)}{[a(1-b)+b(1-a)]^2}. \tag{3.2}$$

Furthermore, we have

$$\frac{\partial^2 P}{\partial a^2} = \frac{-2b(1-b)(1-2b)}{[a(1-b)+b(1-a)]^3},$$

so that, as a function of a, it follows that $P(a,b)$ is concave up for $\frac{1}{2} < b < 1$ and concave down for $0 < b < \frac{1}{2}$. Similarly,

$$\frac{\partial^2 P}{\partial b^2} = \frac{2a(1-a)(1-2a)}{[a(1-b)+b(1-a)]^3}.$$

Adams makes an interesting remark relating to the mixed second partial derivative

$$\frac{\partial^2 P}{\partial a \partial b} = \frac{a-b}{[a(1-b)+b(1-a)]^3}. \tag{3.3}$$

It follows from (3.3) that $\frac{\partial P}{\partial a}$, viewed as a function of b, is increasing for $b < a$ and decreasing for $b > a$, so it is maximized as a function of b when $b = a$. Since $\frac{\partial P}{\partial a}$ is positive for every $0 < b < 1$, it must be most positive when $b = a$. Alternatively, (3.3) tells us that $\frac{\partial P}{\partial b}$, viewed as a function of a, is increasing for $a > b$ and decreasing for $a < b$, so it is minimized as a function of a when $a = b$. Since $\frac{\partial P}{\partial b}$ is negative for every $0 < a < 1$, we conclude that it is most negative when $a = b$.

Adams interprets these facts in the following manner: since $P(a,b)$ increases most rapidly with a when $b = a$ (and decreases most rapidly with b when $a = b$), one should field one's strongest team when playing an opponent of equal strength [8]. Once again, this observation is perhaps more interesting in sports other than baseball, where the star players (other than pitchers) play nearly every game when healthy, although James points out that Yankees manager Casey Stengel preferred to save his ace pitcher, Whitey Ford, for the strongest opposition. It seems particularly relevant to European soccer, where the best teams engage in several different competitions at the same time against opponents of varying quality, and even the top players must occasionally be rested.

In principle, there are two ways to increase the value of $P(a,b)$: by increasing a or by decreasing b. Under most circumstances, a team can only control its own quality and not that of its opponent. There are some situations, however, such as the Yankees signing a key player away from the Red Sox, where an individual or entity might exercise a degree of control over both teams. Similarly, there are many two-player games (such as Parcheesi and backgammon) in which each player's move affects the position of both players. In any such setting, it is a legitimate question whether the priority of an individual or team should be to improve its own standing or to diminish that of its adversary.

Recall that the gradient of a function signifies the direction of the greatest rate of increase. The next result, which has apparently escaped notice until now, follows directly from equations (3.1) and (3.2).

Proposition 3. *For any point (a, b), except where a and b both belong to the set $\{0, 1\}$, the gradient of the James function $P(a, b)$ is a positive multiple of the vector*

$$\langle b(1-b), -a(1-a)\rangle.$$

In other words, to maximize the increase $P(a, b)$, the optimal ratio of the increase of a to the decrease of b is $b(1-b) : a(1-a)$.

One consequence of this result is that when two teams have identical winning percentages, the optimal strategy for increasing $P(a, b)$ is to increase a and to decrease b in equal measure. The same fact holds when two teams have complementary winning percentages. In all other situations, the maximal increase of $P(a, b)$ is achieved by increasing a and decreasing b by different amounts, with the ratio tilted towards the team whose winning percentage is further away from $\frac{1}{2}$. In the extremal cases, when one of the two values a or b belongs to the set $\{0, 1\}$, the optimal strategy is to devote all resources to changing the winning percentage of the team that is either perfectly good or perfectly bad. This observation is somewhat vacuous when $a = 1$ or $b = 0$, since $P(a, b)$ is already as large as it could possibly be, although the strategy is entirely reasonable when $a = 0$ or $b = 1$. It also makes sense that the gradient is undefined at the points $(0, 0)$, $(0, 1)$, $(1, 0)$, and $(1, 1)$, since these winning percentages do not provide enough information to determine how much one team must improve to defeat the other.

If $P(a, b) = c$, it is easy to see that $a(1-b)(1-c) = (1-a)bc$, which implies the next result.

Proposition 4. *If $0 < a < 1$, then $P(a, b) = c$ if and only if $P(a, c) = b$. In other words, for a fixed value of a, the James function is an involution.*

The practical interpretation of this result is simple to state, even if it is not intuitively obvious: if team A has probability c of beating a team with winning percentage b, then team A has probability b of beating a team with winning percentage c. The James conditions already imply this relationship whenever b and c both belong to the set $\{0, 1\}$ or the set $\{\frac{1}{2}, a\}$. Nevertheless, it is not evident at this point whether the involutive property is a necessary consequence of the James conditions. (Example 7 will provide an answer to this question.)

Proposition 4 has two further implications that are worth mentioning. The first is a variation of the involutive property that holds for a fixed value of b:

If $0 < b < 1$, then $P(a, b) = c$ if and only if $P(1-c, b) = 1-a$.

The second is that the level curves for the James function (that is, the set of all points for which $P(a, b) = c$ for a particular constant c) can be written

$$b = P(a, c) = \frac{a(1-c)}{a(1-c) + c(1-a)} \tag{3.4}$$

for $0 < a < 1$. These level curves are the concrete manifestation of a straightforward principle: if a team A improves by a certain amount, there should be a corresponding amount that a team B can improve so that the probability of A defeating B remains unchanged. Each level curve represents the path from $(0, 0)$ to $(1, 1)$ that such a pair would take in tandem.

We conclude this section with one more observation relating to these level curves.

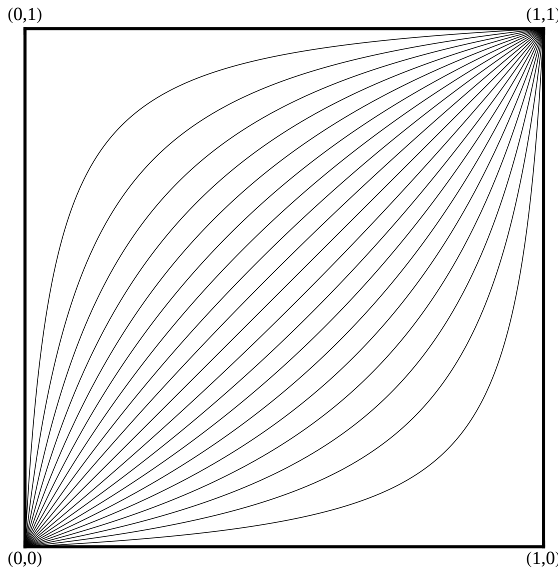

FIGURE 1. The level curves for the James function $P(a,b)$.

Proposition 5. *For any $0 < c < 1$, the corresponding level curve for the James function $P(a,b)$ is the unique solution to the differential equation*

$$\frac{db}{da} = \frac{b(1-b)}{a(1-a)}$$

that passes through the point $(c, \frac{1}{2})$.

Another way of stating this result is that, for two teams to maintain the same value of $P(a,b)$, they should increase (or decrease) their winning percentages according to the ratio $a(1-a) : b(1-b)$. One can either verify this assertion directly, by solving the differential equation to obtain (3.4), or by appealing to Proposition 3 and recalling that the gradient is always perpendicular to the level curve at a particular point.

4. Jamesian functions

We will now consider the question of whether there is a unique function satisfying the James conditions. We begin with the following observation, which is implicit in the construction of the James function.

Proposition 6. *The James function is the only function derived from the Bradley–Terry model that satisfies the James conditions (or the proto-James conditions).*

Proof. Suppose that $\pi(A,B)$ is such a function. Let A be a team with arbitrary winning percentage $0 < a < 1$ and let C be a team with winning percentage $\frac{1}{2}$. Condition (a), or alternatively condition (2), implies that

$$a = \pi(A,C) = \frac{w(A)}{w(A)+w(C)}.$$

Solving for $w(A)$, we obtain

$$w(A) = \frac{aw(C)}{1-a} = q_c(a),$$

where $c = w(C)$. Thus $\pi(A, B)$ is identical to the James function $P(a, b)$. □

Let S denote the open unit square $(0,1) \times (0,1)$. We will say that any function $J(a, b)$, defined on the set $\overline{S} \setminus \{(0,0) \cup (1,1)\}$, that satisfies the James conditions is a *Jamesian function*. Our immediate objective is to disprove Conjecture 1 by identifying at least one example of a Jamesian function that is different from the James function $P(a, b)$. Proposition 6 guarantees that any such function, if it exists, cannot be derived from the Bradley–Terry model.

Example 7. We will reverse-engineer our first example of a new Jamesian function by starting with its level curves. Consider the family of curves $\{j_c\}_{c\in(0,1)}$ defined as follows:

$$j_c(a) = \begin{cases} \dfrac{a}{2c}, & 0 < a \le \dfrac{2c}{1+2c} \\ 2ca + 1 - 2c, & \dfrac{2c}{1+2c} < a < 1 \end{cases}$$

for $0 < c \le \frac{1}{2}$ and

$$j_c(a) = \begin{cases} (2-2c)a, & 0 < a \le \dfrac{1}{3-2c} \\ \dfrac{a+1-2c}{2-2c}, & \dfrac{1}{3-2c} < a < 1 \end{cases}$$

for $\frac{1}{2} < c < 1$. These curves have been chosen to satisfy certain symmetry properties, which the reader can probably deduce but which we will not state explicitly. We define the function $J(a, b)$ on S by assigning to every point (a, b) the value of c associated with the particular curve j_c that passes through that point. We assign the value 0 or 1 to points on the boundary of S, as dictated by the James conditions.

A bit more work yields an explicit formula for $J(a, b)$, from which one can verify directly that all of the James conditions are satisfied:

$$J(a,b) = \begin{cases} \dfrac{a}{2b}, & (a,b) \in \mathrm{I} \\[2ex] \dfrac{2a-b}{2a}, & (a,b) \in \mathrm{II} \\[2ex] \dfrac{1-b}{2(1-a)}, & (a,b) \in \mathrm{III} \\[2ex] \dfrac{1+a-2b}{2(1-b)}, & (a,b) \in \mathrm{IV} \end{cases},$$

where I, II, III, and IV are subsets of $\overline{S} \setminus \{(0,0) \cup (1,1)\}$ that are defined according to Figure 3.

Observe that the appropriate definitions coincide on the boundaries between regions, from which it follows that $J(a, b)$ is continuous on $\overline{S} \setminus \{(0,0) \cup (1,1)\}$. On the other hand, it is not difficult to see that $J(a, b)$ fails to be differentiable at all points of the form $(a, 1-a)$ for $0 < a < \frac{1}{2}$ or $\frac{1}{2} < a < 1$. (With some effort, one

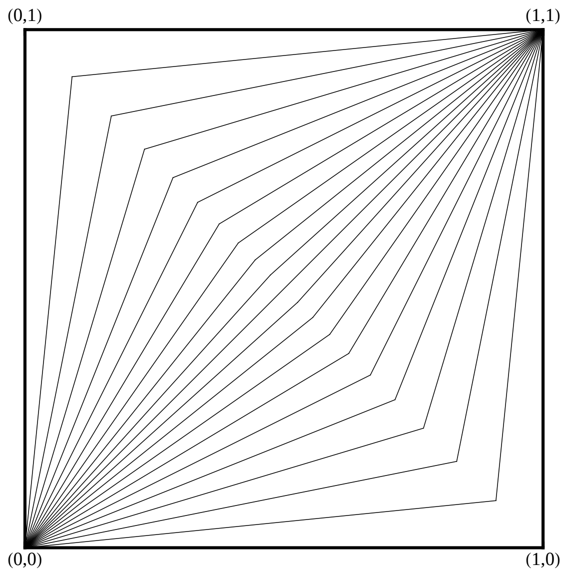

FIGURE 2. The level curves for the function $J(a, b)$ in Example 7.

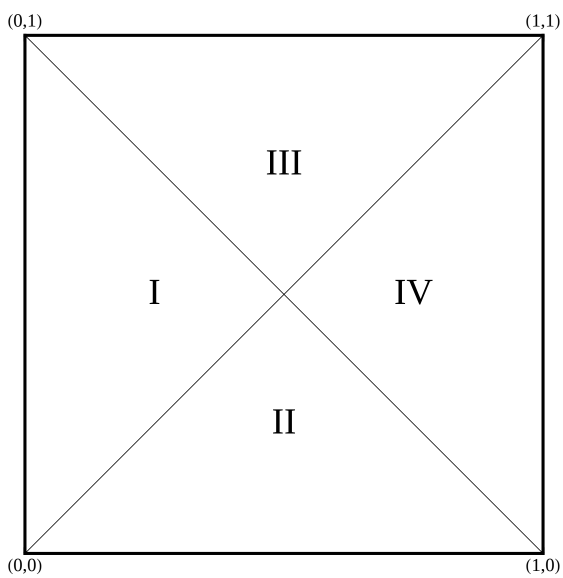

FIGURE 3. The subsets of $\overline{S} \setminus \{(0,0) \cup (1,1)\}$ in Example 7.

can show that it is differentiable at the point $(\frac{1}{2}, \frac{1}{2})$.) In reference to Proposition 4, note that $J(\frac{1}{3}, \frac{1}{4}) = \frac{5}{8}$ and $J(\frac{1}{3}, \frac{5}{8}) = \frac{4}{15}$. In other words, the involutive property is not a necessary consequence of the James conditions.

In view of the preceding example, we need to refine our terminology somewhat. We will refer to any Jamesian function (such as the James function itself) that satisfies the condition

$$J\big(a, J(a,b)\big) = b$$

for $0 < a < 1$ as an *involutive Jamesian function*. We will encounter many such functions in Section 5.

It is remarkably easy to construct Jamesian functions with discontinuities in S, as we will see in our next example.

Example 8. Let $J(a,b)$ be any Jamesian function and take $0 < \gamma < 1$. Define the function $J(a,b;\gamma)$ as follows:

$$J(a,b;\gamma) = \begin{cases} J(a,b), & (a,b) \in A = \big([0,\frac{1}{2}] \times [0,\frac{1}{2}]\big) \setminus \{(0,0)\} \\ \gamma J(a,b), & (a,b) \in B = [0,\frac{1}{2}) \times (\frac{1}{2},1] \\ \gamma J(a,b) + 1 - \gamma, & (a,b) \in C = (\frac{1}{2},1] \times [0,\frac{1}{2}) \\ J(a,b), & (a,b) \in D = \big([\frac{1}{2},1] \times [\frac{1}{2},1]\big) \setminus \{(1,1)\} \end{cases},$$

as illustrated in Figure 4.

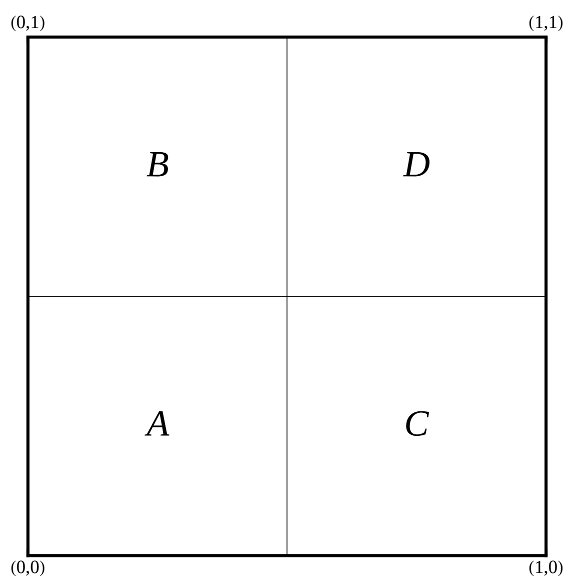

FIGURE 4. The subsets of $\overline{S} \setminus \{(0,0) \cup (1,1)\}$ in Example 8.

It should be apparent from the definition that $J(a,b;\gamma)$ satisfies James conditions (a) and (b). Condition (e) follows from the inequalities

$$0 \le \gamma x \le x$$

and

$$x \le \gamma x + 1 - \gamma \le 1$$

for $0 \le x \le 1$, along with the fact that γx and $\gamma x + 1 - \gamma$ are both strictly increasing functions of x. We will now verify the two remaining conditions:

(c) If a point (a, b) belongs either to the subset A or the subset D, then (b, a) belongs to the same subset. In either case,

$$J(b, a; \gamma) = J(b, a) = 1 - J(a, b) = 1 - J(a, b; \gamma).$$

If (a, b) belongs to B, then (b, a) belongs to C, from which it follows that

$$\begin{aligned} J(b, a; \gamma) = \gamma J(b, a) + 1 - \gamma =& \gamma\big(1 - J(a, b)\big) + 1 - \gamma \\ =& 1 - \gamma J(a, b) = 1 - J(a, b; \gamma). \end{aligned}$$

An analogous argument pertains to the case where (a, b) belongs to C.

(d) If a point (a, b) belongs either to B or to C, then $(1 - b, 1 - a)$ belongs to the same subset. If (a, b) belongs to A or D, then $(1 - b, 1 - a)$ belongs to D or A respectively. In all these cases, it follows that $J(1 - b, 1 - a; \gamma) = J(a, b; \gamma)$, since $J(a, b)$ possesses the same property.

In other words, every function defined in this manner for $0 < \gamma < 1$ must be a Jamesian function. Note that $J(a, b; \gamma)$ is discontinuous at every point of the form $(a, \frac{1}{2})$ for $0 < a < 1$ or $(\frac{1}{2}, b)$ for $0 < b < 1$, which is the purpose of this example. If we extend this construction to $\gamma = 1$, we simply obtain the original Jamesian function $J(a, b)$. If we take $\gamma = 0$, the resulting function satisfies the proto-James conditions but not the James conditions.

Returning to the situation where $0 < \gamma < 1$, we remark that $J(a, b; \gamma)$ cannot be involutive if $J(a, b)$ is involutive. Suppose that $J(\frac{1}{3}, \frac{1}{4}) = c$ and $J(\frac{1}{3}, c) = \frac{1}{4}$. Condition (e) implies that

$$c = J(\tfrac{1}{3}, \tfrac{1}{4}) > J(\tfrac{1}{4}, \tfrac{1}{4}) = \tfrac{1}{2},$$

so the point $(\frac{1}{3}, c)$ belongs to the subset B, whereas $(\frac{1}{3}, \frac{1}{4})$ belongs to A. Consequently

$$J(\tfrac{1}{3}, \tfrac{1}{4}; \gamma) = J(\tfrac{1}{3}, \tfrac{1}{4}) = c$$

but

$$J(\tfrac{1}{3}, c; \gamma) = \gamma J(\tfrac{1}{3}, c) < \tfrac{1}{4}.$$

Another way of stating this observation is that, for a particular Jamesian function $J(a, b)$, the function $J(a, b; \gamma)$ can be involutive for no more than one value $0 < \gamma \leq 1$.

Our next proposition, whose proof we leave to the reader, is similar in spirit to the preceding example.

Proposition 9. *Suppose that $J_1(a, b), J_2(a, b), \ldots, J_n(a, b)$ are Jamesian functions and let $\alpha_1, \alpha_2, \ldots, \alpha_n$ be positive constants with $\alpha_1 + \alpha_2 + \ldots + \alpha_n = 1$. The linear combination*

$$\alpha_1 J_1(a, b) + \alpha_2 J_2(a, b) + \ldots + \alpha_n J_n(a, b)$$

is also a Jamesian function.

Having considered some of the more pathological examples of Jamesian functions, we will devote the next section to constructing functions that are involutive, continuous, and (in many cases) differentiable.

5. Hyper-James functions

Let $g : (0,1) \to \mathbb{R}$ be a continuous, strictly increasing function that satisfies the conditions

- $g(1-a) = -g(a)$.
- $\lim_{a\to 0^+} g(a) = -\infty$.

These conditions imply that $g(\frac{1}{2}) = 0$ and that

$$\lim_{a\to 1^-} g(a) = \infty.$$

Observe that $g^{-1} : \mathbb{R} \to (0,1)$ is a continuous, strictly increasing function with $g^{-1}(-a) = 1 - g^{-1}(a)$. It makes sense to define $g(0) = -\infty$ and $g(1) = \infty$, so that $g^{-1}(-\infty) = 0$ and $g^{-1}(\infty) = 1$. We claim that any such function g can be used to construct a Jamesian function.

Theorem 10. *For any g satisfying the conditions specified above, the function*

$$J(a,b) = g^{-1}\big(g(a) - g(b)\big) \tag{5.1}$$

is a continuous involutive Jamesian function on $\overline{S} \setminus \{(0,0) \cup (1,1)\}$.

Proof. Consider each of the five James conditions:
(a) Note that

$$J(a, \tfrac{1}{2}) = g^{-1}\big(g(a) - g(\tfrac{1}{2})\big) = g^{-1}\big(g(a) - 0\big) = a.$$

(b) Similarly,

$$J(a,0) = g^{-1}\big(g(a) - g(0)\big) = g^{-1}(g(a) + \infty) = 1$$

for $0 < a \le 1$. (The case where $a = 0$ yields the indeterminate form $-\infty + \infty$.)
(c) Observe that

$$J(b,a) = g^{-1}\big(g(b) - g(a)\big) = 1 - g^{-1}\big(g(a) - g(b)\big) = 1 - J(a,b).$$

(d) Likewise,

$$J(1-b, 1-a) = g^{-1}\big(g(1-b) - g(1-a)\big) = g^{-1}\big(-g(b) + g(a)\big) = J(a,b).$$

(e) The fact that $J(a,b)$ is a strictly increasing function of a follows from the fact that both g and g^{-1} are strictly increasing.

It is not difficult to see that $J(a,b)$ is involutive:

$$\begin{aligned} J\big(a, J(a,b)\big) &= g^{-1}\Big(g(a) - g\big(g^{-1}\big(g(a) - g(b)\big)\big)\Big) \\ &= g^{-1}\big(g(a) - g(a) + g(b)\big) \\ &= g^{-1}\big(g(b)\big) = b, \end{aligned}$$

as long as $0 < a < 1$. The continuity of $J(a,b)$ is an immediate consequence of the continuity of g and g^{-1}. □

It is easy to use this theorem to generate concrete examples.

Example 11. The function

$$g(a) = \frac{2a-1}{a(1-a)}$$

satisfies all the necessary conditions for Theorem 10, so (5.1) defines an involutive Jamesian function that is continuous on $\overline{S} \setminus \{(0,0) \cup (1,1)\}$. Since

$$g^{-1}(s) = \frac{s - 2 + \sqrt{s^2+4}}{2s},$$

we obtain

$$J(a,b) = \frac{x + y - \sqrt{x^2+y^2}}{2y} = \frac{x}{x + y + \sqrt{x^2+y^2}},$$

where $x = 2ab(1-a)(1-b)$ and $y = (b-a)(2ab - a - b + 1)$.

Example 12. The function $g(a) = -\cot(\pi a)$ yields the continuous involutive Jamesian function

$$J(a,b) = \frac{1}{\pi}\cot^{-1}\big(\cot(\pi a) - \cot(\pi b)\big),$$

where we are using the version of the inverse cotangent that attains values between 0 and π.

The construction described in Theorem 10 is closely related to what is known as a *linear model* for paired comparisons. In such a model,

$$\pi(A,B) = H\big(v(A) - v(B)\big),$$

where v denotes a measure of worth and H is the cumulative distribution function of a random variable that is symmetrically distributed about 0 (see [2, Section 1.3]). The Bradley–Terry model can be viewed as a linear model, where H is the logistic function

$$H(s) = \frac{e^s}{e^s+1} = \int_{-\infty}^{s} \frac{e^t}{(1+e^t)^2}\,dt$$

and $v(A) = \log w(A)$. In particular, the James function can be constructed in the manner of Theorem 10, with $H = g^{-1}$ being the logistic function and g being the so-called logit function

$$g(a) = \log\left(\frac{a}{1-a}\right).$$

(This observation could charitably be construed as an *a posteriori* justification for the term "log5" originally used by James.)

What is distinctive about the James function in this context is that the construction is symmetric, with $v(A) = \log w(A)$ and $v(B) = \log w(B)$ replaced by $g(a) = \log(a/(1-a))$ and $g(b) = \log(b/(1-b))$ respectively. This symmetry corresponds to the twofold application of the Bradley–Terry model that was discussed in Section 1. Likewise, the fact that both g and g^{-1} appear in the general formulation of Theorem 10 can be interpreted as a consequence of the same model being used to define both worth and probability.

Example 13. Take

$$H(s) = g^{-1}(s) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{s} e^{-\frac{t^2}{2}}\,dt,$$

so that g is the so-called probit function. The involutive Jamesian function $J(a,b) = g^{-1}\big(g(a) - g(b)\big)$ can be considered the analogue of the James function relative to the Thurstone–Mosteller model (see [2]).

Theorem 10 allows us to identify a large class of functions that can be viewed as generalizations of the James function. Since

$$\log\left(\frac{a}{1-a}\right) = \int_{\frac{1}{2}}^{a} \left(\frac{1}{t} + \frac{1}{1-t}\right) dt = \int_{\frac{1}{2}}^{a} \frac{1}{t(1-t)} dt,$$

we define

$$g_n(a) = \int_{\frac{1}{2}}^{a} \frac{1}{(t(1-t))^n} dt \tag{5.2}$$

for $n \geq 1$. It is not difficult to verify that g_n satisfies all of the prescribed requirements for Theorem 10. (The stipulation that $g_n(0) = -\infty$ precludes the case where $0 < n < 1$.) Define

$$P_n(a,b) = g_n^{-1}\big(g_n(a) - g_n(b)\big). \tag{5.3}$$

For $n > 1$, we shall refer to $P_n(a,b)$ as a *hyper-James function*. Each of these functions is an involutive Jamesian function and is continuous on the set $\overline{S} \setminus \{(0,0) \cup (1,1)\}$.

In some situations, such as when n is a positive integer, it is possible to obtain a more concrete representation for the function g_n. We require the following partial fractions expansion, which is easier to prove than it is to guess: if m and n are positive integers, then

$$\frac{1}{t^m(1-t)^n} = \sum_{k=0}^{m-1} \binom{n+k-1}{k} \frac{1}{t^{m-k}} + \sum_{k=0}^{n-1} \binom{m+k-1}{k} \frac{1}{(1-t)^{n-k}}. \tag{5.4}$$

(Euler knew a similar formula no later than 1771; see [3, pp. 222–23].) In view of this observation, it is straightforward to show that (5.2) can be rewritten

$$g_n(a) = \binom{2n-2}{n-1} \log\left(\frac{a}{1-a}\right) + \sum_{k=0}^{n-2} \binom{n+k-1}{k} \frac{1}{n-k-1} \left(\frac{1}{(1-a)^{n-k-1}} - \frac{1}{a^{n-k-1}}\right)$$

for any integer $n \geq 1$, where the sum is empty if $n = 1$. While this formula is certainly worth having, we still cannot obtain an explicit representation for $P_n(a,b)$ without a similar formula for g_n^{-1}.

It is also possible to evaluate (5.2) with n replaced by $m + \frac{1}{2}$ for a positive integer m. Substituting $t = \sin^2\theta$, we have

$$\int_{\frac{1}{2}}^{a} \frac{dt}{(t(1-t))^{m+\frac{1}{2}}} = 2\int_{\frac{\pi}{4}}^{\sin^{-1}\sqrt{a}} \frac{d\theta}{\sin^{2m}\theta \cos^{2m}\theta}.$$

One can break this expression apart using equation (5.4), with $t = \sin^2\theta$, but it is better to write

$$\begin{aligned} 2\int_{\frac{\pi}{4}}^{\sin^{-1}\sqrt{a}} \frac{d\theta}{\sin^{2m}\theta \cos^{2m}\theta} &= 2^{2m} \int_{\frac{\pi}{4}}^{\sin^{-1}\sqrt{a}} \frac{2\,d\theta}{\sin^{2m}(2\theta)} \\ &= 2^{2m} \int_{\frac{\pi}{2}}^{2\sin^{-1}\sqrt{a}} \left(\cot^2\phi + 1\right)^{m-1} \csc^2\phi\, d\phi. \end{aligned}$$

Substituting $u = \cot\phi$ gives us

$$\begin{aligned} g_{m+\frac{1}{2}}(a) =& 2^{2m} \int_{\frac{1-2a}{2\sqrt{a(1-a)}}}^{0} (u^2+1)^{m-1}\, du \\ =& \sum_{k=0}^{m-1} \binom{m-1}{k} \frac{2^{2m-2k-1}}{2k+1} \left(\frac{2a-1}{\sqrt{a(1-a)}} \right)^{2k+1}. \end{aligned}$$

In particular, taking $m = 1$ we have

$$g_{\frac{3}{2}}(a) = \frac{2(2a-1)}{\sqrt{a(1-a)}}.$$

In this case, since

$$g_{\frac{3}{2}}^{-1}(s) = \frac{s + \sqrt{s^2+16}}{2\sqrt{s^2+16}},$$

we see that

$$P_{\frac{3}{2}}(a,b) = \frac{1}{2} + \frac{v'\sqrt{u} - u'\sqrt{v}}{2\sqrt{u + v - 4uv - 2u'v'\sqrt{uv}}},$$

where $u = a(1-a)$, $v = b(1-b)$, $u' = 1-2a$, and $v' = 1-2b$.

In general, it seems unlikely that there is an explicit formula for $P_n(a,b)$ that is more useful than equation (5.3).

We will now turn our attention to the issue of differentiability. For any function defined according to Theorem 10, a routine calculation shows that

$$\frac{\partial J}{\partial a} = \frac{g'(a)}{g'\big(J(a,b)\big)} \tag{5.5}$$

and

$$\frac{\partial J}{\partial b} = \frac{-g'(b)}{g'\big(J(a,b)\big)} \tag{5.6}$$

at all points (a,b) for which the above quotients are defined. Based on this observation, we are able to obtain the following result.

Proposition 14. *If g is continuously differentiable on $(0,1)$, with g' never equal to 0, the corresponding Jamesian function $J(a,b)$ is differentiable on S. Conversely, if $J(a,b)$ is differentiable on S, the function g must be differentiable on $(0,1)$ with g' never 0.*

Proof. Suppose that g' is continuous and nonzero on $(0,1)$. It follows from (5.5) and (5.6) that both $\frac{\partial J}{\partial a}$ and $\frac{\partial J}{\partial b}$ are defined and continuous at all points in the open set S, which guarantees that $J(a,b)$ is differentiable on S.

Now suppose that $J(a,b)$ is differentiable at every point in S. Let a_0 be an arbitrary element of $(0,1)$. Since g is strictly increasing, it could only fail to be differentiable on a set of measure 0 (see [13, p. 112]). In particular, there is at least one c in $(0,1)$ for which $g'(c)$ is defined. Since $J(a_0,b)$, viewed as a function of b, attains every value in the interval $(0,1)$, there exists a b_0 in $(0,1)$ such that $J(a_0,b_0) = c$. Note that

$$g(a) = g\big(J(a,b_0)\big) + g(b_0)$$

for all a in $(0,1)$, so the chain rule dictates that

$$g'(a_0) = g'(c) \cdot \frac{\partial J}{\partial a}(a_0, b_0).$$

Therefore g is differentiable on the entire interval $(0,1)$. Suppose, for the sake of contradiction, that there were some d in $(0,1)$ for which $g'(d) = 0$. As before, there would exist a b_1 in $(0,1)$ such that $J(a_0, b_1) = d$, which would imply that

$$g'(a_0) = g'(d) \cdot \frac{\partial J}{\partial a}(a_0, b_1) = 0.$$

Consequently g' would be identically 0 on $(0,1)$, which is impossible. □

In other words, all the specific examples of Jamesian functions we have introduced in this section, including the hyper-James functions, are differentiable on S. We can now state a more general version of Proposition 3, which follows directly from (5.5) and (5.6).

Proposition 15. *For any differentiable Jamesian function $J(a,b)$ defined according to Theorem 10, the gradient at a point (a,b) in S is a positive multiple of the vector $\langle g'(a), -g'(b)\rangle$.*

If g is differentiable on $(0,1)$, the condition that $g(1-a) = -g(a)$ implies that $g'(1-a) = g'(a)$. Hence the gradient of $J(a,b)$ is a positive multiple of $\langle 1, -1\rangle$ whenever $b = a$ or $b = 1-a$. This observation generalizes the fact that, whenever two teams have identical or complementary winning percentages, the optimal strategy for increasing $P(a,b)$ is to increase a and decrease b by equal amounts.

Since any Jamesian function given by (5.1) is involutive, the level curve $J(a,b) = c$ for $0 < c < 1$ can be rewritten

$$b = J(a,c) = g^{-1}\big(g(a) - g(c)\big),$$

or $g(a) = g(b) + g(c)$. Note that each curve is a strictly increasing function of a that connects the points $(0,0)$ and $(1,1)$. Moreover, we have the following generalization of Proposition 5.

Proposition 16. *Let $J(a,b)$ be a differentiable Jamesian function defined according to Theorem 10. For any $0 < c < 1$, the corresponding level curve for $J(a,b)$ is the unique solution to the differential equation*

$$\frac{db}{da} = \frac{g'(a)}{g'(b)}$$

that passes through the point $(c, \frac{1}{2})$.

Thus the level curves for the Jamesian functions defined in Examples 11 and 12 are given by the differential equations

$$\frac{db}{da} = \frac{(2a^2 - 2a + 1)(b(1-b))^2}{(2b^2 - 2b + 1)(a(1-a))^2}$$

and

$$\frac{db}{da} = \left(\frac{\sin(\pi b)}{\sin(\pi a)}\right)^2$$

respectively. Likewise, the level curves for any hyper-James function $P_n(a,b)$ are given by the differential equation

$$\frac{db}{da} = \left(\frac{b(1-b)}{a(1-a)}\right)^n .$$

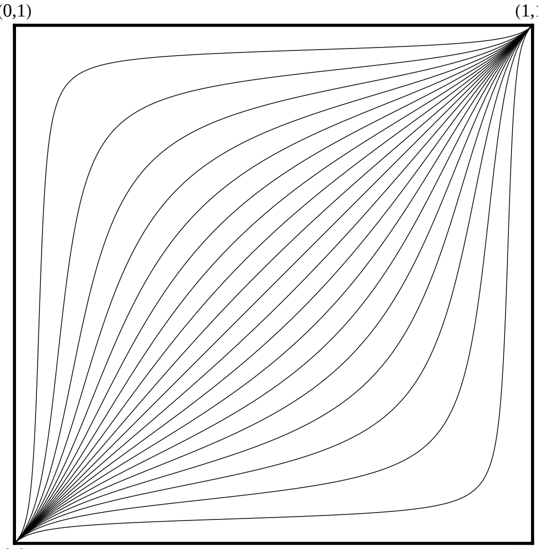

Figure 5. The level curves for the hyper-James function $P_2(a,b)$.

Acknowledgments

We would never have written this paper if Caleb Garza, an undergraduate at Connecticut College, had not decided to give a senior seminar talk on a topic from sabermetrics. We are sincerely grateful to him for prompting (or reviving) our interest in this material and for bringing the work of the third-named author to the attention of the first two.

References

[1] Ralph Allan Bradley and Milton E. Terry, Rank analysis of incomplete block designs. I. The method of paired comparisons, *Biometrika* **39** (1952), 324–345.
[2] H. A. David, *The Method of Paired Comparisons*, 2nd ed., Oxford University Press, New York, 1988.
[3] Leonhard Euler, Meditationes circa singulare serierum genus, *Opera Omnia*, Series Prima, Volume XV, edited by Georg Faber, B. G. Teubner, Leipzig, 1927.
[4] Mark E. Glickman, Introductory note to 1928 (= 1929), in *Ernst Zermelo: Collected Works*, Volume II, edited by Heinz-Dieter Ebbinghaus and Akihiro Kanamori, Springer-Verlag, Berlin, 2013, pp. 616–621.
[5] Mark E. Glickman and Albyn C. Jones, Rating the chess rating system, *Chance* **12** (1999), no. 2, 21–28.
[6] I. J. Good, On the marking of chess-players, *Math. Gaz.* **39** (1955), 292–296.
[7] Bill James, *1980 Baseball Abstract*, self-published, Lawrence, KS, 1980.
[8] Bill James, *1981 Baseball Abstract*, self-published, Lawrence, KS, 1981.

[9] T. Jech, A quantitative theory of preferences: some results on transition functions, *Soc. Choice Welf.* **6** (1989), no. 4, 301–314.

[10] R. Duncan Luce, *Individual Choice Behavior: A Theoretical Analysis*, John Wiley and Sons, New York, 1959.

[11] Steven J. Miller, A justification of the log5 rule for winning percentages, `http://web.williams.edu/Mathematics/sjmiller/public_html/103/Log5WonLoss_Paper.pdf`, 2008.

[12] Steven J. Miller, Taylor Corcoran, Jennifer Gossels, Victor Luo, and Jaclyn Porfilio, Pythagoras at the bat, in *Social Networks and the Economics of Sports*, edited by Victor Zamaraev, Springer-Verlag, Berlin, to appear.

[13] H. L. Royden and P. M. Fitzpatrick, *Real Analysis*, 4th ed., Prentice Hall, Boston, 2010.

[14] Michael Stob, A supplement to "A mathematician's guide to popular sports", *Amer. Math. Monthly* **91** (1984), no. 5, 277–282.

[15] E. Zermelo, Die Berechnung der Turnier-Ergebnisse als ein Maximumproblem der Wahrscheinlichkeitsrechnung, *Math. Z.* **29** (1929), no. 1, 436–460; The calculation of the results of a tournament as a maximum problem in the calculus of probabilities, in *Ernst Zermelo: Collected Works*, Volume II, edited by Heinz-Dieter Ebbinghaus and Akihiro Kanamori, Springer-Verlag, Berlin, 2013, pp. 622–671.

Department of Mathematics, Connecticut College, New London, CT 06320
E-mail address: `cnham@conncoll.edu`
E-mail address: `wpjoh@conncoll.edu`

Department of Mathematics and Statistics, Williams College, Williamstown, MA 01267
E-mail address: `sjm1@williams.edu`