Performance Tournaments with Crowdsourced Judges

Daryl Pregibon, William D Heavlin, Google, Inc.

August 23, 2013

Abstract

A performance slam is a competition among a fixed set of performances whereby pairs of performances are judged by audience participants. When performances are recorded on electronic media, performance slams become amenable to audiences that watch online and judge asynchronously ("crowdsourced"). In order to better entertain the audience, we want to show the better performances ("exploitation"). In order to identify the good videos, we want to glean a least some information about all videos ("exploration").

Our approach has three elements: (1) We take our preference model from Bradley and Terry (1952). (2) Its parameters we calculate by rewriting the likelihood gradient into a fixed point estimate, one which mimics the estimate of Mantel and Haenszel (1959). (3) Each pair of performances is chosen sequentially, always chosen to minimize the weighted variance of (the logarithms of) the Bradley-Terry parameter estimates. Our preferred weights consist of the logrank weights proposed by Savage (1956).

Key words: Bradley-Terry model, exploration and exploitation, mojovian, optimal ranking, preference modeling, Savage scores.

1. Background

1.1 Slams on YouTube

YouTube is a popular website devoted to sharing videos. The repository consists of billions of videos of widely varying quality and mass appeal. It is said that Justin Beeber got his start to fame through YouTube videos supplied by his family. Recently, Gangham Style demonstrated the viral nature and mass appeal of video content. Identifying such content among the millions of new videos entering the repository each week is an important challenge and key to the continued success of YouTube. This paper focuses on the latter, and in particular, on one approach to measuring the quality of the user experience.

A *slam* consists of a certain kind of competition in which the audience rates a series of performances; like most competitions, the goal is to determine the winner, and more generally to rank the competitors. Marc Smith is credited with the first slam, for poetry, in Chicago, in 1984. Over time, the slam format has generally converged to pairs of performances: Cognitively, slams of size two are easier to compare, tasking as they do the audience members' recollections less. In cultural assumptions and by data format, slams of size two align with sports competitions, with each round resulting in a winner and a loser.

In late 2011 into early 2012, YouTube ran a series of slam tournaments within five genre: comedy, music, dance, cute, and one other ("bizarre"). Viewers would watch a pair of videos, and those who had signed into YouTube could vote for the one they liked better.

Note that the slam format on YouTube differs from slams with live audiences in at least four ways: First, YouTube performances are video recordings, so a second or third performance by the same artist is always identical to the first ("perfect replicability"). Second, any given pair of performances is always seen by only one user ("audiences of size one"), so each slam receives only one vote. Third, in the YouTube format, the decision of a winner can be delayed, since there is no mutually visible collective awaiting for the final outcome ("delayed tournament outcome"). Fourth, in common with other electronic media, voting on YouTube has the potential of being undesirably manipulated by automatic voting programs ("the spam issue").

In spite of such differences, for YouTube the two central challenges of the slam format remain: (1) Someone has to pair up some performances to create the slams ("the scheduling problem"). (2) Somehow one has to determine the winner(s) and/or identify the better ones ("the ranking problem").

1.2 Models of Pairwise Competition

When Thurstone (1927) introduced his law of comparative judgment, he offered three ideas: (a) Humans find the task of judging pairs of objects easier than directly ranking longer lists. (b) One can recover the underlying ranking by modeling a single underlying interval scale $\mu(A)$; the probability of preferring object A over object B is a function of the arithmetic difference$\mu(A) - \mu(B)$. (c) The probability can be modeled by the CDF of any symmetric distribution, and Thurstone picked the fixed-variance Gaussian. Turner and Firth (2012) implement Thurstone's model in the CRAN package BradleyTerry2.

Bradley and Terry (1952) developed a model similar to Thurstone's, replacing the Gaussian CDF with that of the CDF associated with continuous logistic. In their model, the difference $\mu(A) - \mu(B)$ acquires the interpretation of the logarithm of a preference odds. We detail this further in section 2. Note that BradleyTerry2 supports this model by offering a logit link function.

Elo's (1978) chess rating system is essentially a linearized Bradley-Terry model with participants-only updates. Glickman (1993, 1999) modifies Elo's model to include a variance function. Such a variance function recognizes, for example, that players who have not played for a while have greater uncertainty about their true ability. Glickman's algorithms are fundamentally Bayesian, and use MCMC machinery to calculate updates to player strengths.

1.3 Best k of n Selection

The theory of ranking and selection is largely an outgrowth of work by Wald (1939, 1950) and Blackwell and Girshick (1954). Their focus on decision theory provides an alternative formulation to the classical topic of statistical hypothesis testing.

Wetherill and Ofosu (1974) distinguish between minimax, minimax regret, and Bayesian methods. The early literature is dominated by assumptions of normal populations (Gibbons, Olkin, Sobel, 1999). For our purposes, Dunnett's (1960) approach, blending Bayesian and decision theoretic elements, deserves mention: his Bayes risk function is a linear function of the population parameters.

Ranking and selection is intimately tied to sequential statistical procedures. Robbins (1952) introduces the multi-armed bandit problem. Glickman and Jensen (2005) exploit the glicko model framework to propose tournament rounds that maximize Kullback-Liebler distance, and apply it to tournament chess.

1.4 Locally Most Powerful Rank Tests

One class of rank tests involves linear combinations of ranks, for example, to test whether two groups are statistically different. Within this class, there is no uniformly most powerful rank test. However, Lehmann (1959) identifies *locally* most power rank (LMPR) tests. The rationale for the principle of locality is that any increase in statistical power in such hard-to-discriminate circumstances should have reasonably attractive properties when applied to less demanding settings. By LMPR theory, the optimal weights depend on the underlying distribution, f, and are of the form

$$-\frac{\partial \log f(x_{(i)})}{\partial x_{(i)}},$$

where $x_{(i)}$ denotes the i-th order statistic drawn from distribution f.

For example, the van der Waerden weights are asymptotically LMPR-optimal for the normal distribution and Wilcoxon's linear ranks are asymptotically LMPR-optimal for the continuous logistic. For right-skewed distributions, reciprocal ranks are LMPR-optimal for the Pareto distribution and the so-called log-rank or Savage weights are likewise LMPR-optimal for the exponential distribution.

2 Bradley-Terry Model Estimates

2.1 Notation

We have n performances, indexed by $a, b = 1, 2, ..., n$, and T slam pairs, indexed by $t = 1, 2, ..., T$. Each such slam t has two performances, a_t and b_t, and a preference y_t, which takes the value 1 if performance a_t is preferred to b_t, 0 when performance b_t is preferred to a_t, and is otherwise undefined. It is convenient to define the complementary response of $z_t = 1 - y_t$, that is, $z_t = 1$ when b_t is preferred to a_t, 0 when a_t is preferred to b_t, and otherwise undefined. Bradley and Terry (1952) assert this model:

$$Pr\{y_t = 1\} = \frac{\omega(a_t)}{\omega(a_t) + \omega(b_t)}, \tag{1}$$

or, equivalently,

$$\frac{Pr\{y_t = 1\}}{Pr\{y_t = 0\}} = \frac{\omega(a_t)}{\omega(b_t)}. \tag{2}$$

In this notation, each item a has a strength, worth, or *mojo* $\omega(a) > 0$. By equation (2), these parameters are seen to be log-odds, and defined only up to a constant multiplier.

This set notation proves convenient: $a = \{t : a_t = a\} \smile \{t : b_t = a\}$.

2.2 The BT Gradient Estimate

We assume the slams are independent, and index them by t. The likelihood can be written as

$$\ell(\omega) = \prod_t \frac{\omega(a_t)^{y_t}\omega(b_t)^{z_t}}{\omega(a_t) + \omega(b_t)}, \tag{3}$$

and log-likelihood

$$\log\,\ell(\omega) = \sum_t y_t\,\log\,\omega(a_t) + z_t\,\log\,\omega(b_t) - \log\left(\omega(a_t) + \omega(b_t)\right). \tag{4}$$

Differentiating (4) with respect to $\omega(a_t)$ and setting it equal to zero results in this equation,

$$\sum_{t\in a}\left[\frac{y_t}{\omega(a_t)} - \frac{y_t + z_t}{\omega(a_t) + \omega(b_t)}\right] = 0, \tag{5}$$

which Hunter (2004) then solves for $\omega(a)$,

$$\omega(a) = \frac{\sum_{t\in a} y_t}{\sum_{t\in a}\frac{1}{\omega(a)+\omega(b_t)}}, \tag{6}$$

inviting a fixed-point algorithm. Note that equation (5) essentially asserts the total wins for object a equals the expected wins per model (1).

2.3 The BT MH Estimate

In at least one sense, the fixed-point equation (6) is curious, in that it is not obviously symmetric between wins and losses. Indeed, wins by a immediately benefit the numerator of (6), while losses require propagating an update through the network of all other outcomes, and ultimately depend on enforcing the global constraint such as $\prod_a \omega(a) = 1$.These considerations suggest that an improved estimate might be available

Toward this goal. let us define $p(a|a, b) = \omega(a)/(\omega(a) + \omega(b))$. Rewrite the likelihood gradient (5) as

$$0 = \sum_t y_t - (y_t + z_t)\, p(a|a, b_t) = \sum_t y_t\,(1 - p(a|a, b_t)) - z_t\, p(a|a, b_t)$$

$$= \sum_t y_t\, p(b_t|a, b_t) - z_t\, p(a|a, b_t) \tag{7}$$

or, as an equation set to unity,

$$1 = \frac{\sum_t y_t\, p(b_t|a, b_t)}{\sum_t z_t\, p(a|a, b_t)} = \frac{\sum_t y_t\, \omega(b_t)/[\omega(a) + \omega(b_t)]}{\sum_t z_t\, \omega(a)/[\omega(a) + \omega(b_t)]}. \tag{8}$$

The form (8) is recognizably that of the estimate proposed by Mantel and Haenszel (1958). Equation (8) points toward the following fixed-point equations, with superscripts indexing interactions:

$$\omega^0(a) = 1; \upsilon^0(a) = 1$$

$$\upsilon^{k+1}(a) = \frac{\sum_t y_t\, \omega^k(b_t)/[\omega^k(a) + \omega^k(b_t)]}{\sum_t z_t\, \omega^k(a)/[\omega^k(a) + \omega^k(b_t)]}. \tag{9}$$
$$\omega^{k+1}(a) = \omega^k(a)\upsilon^{k+1}(a).$$

Under equations (9), current estimates $\omega^k(a)$ are used to determine residual updates $\upsilon^k(a)$ via a Mantel-Haenszel expression. As an update expression, (9) is visibly symmetric between wins and losses, and empirically we observe that it converges rather quickly.

2.4 Regularization by Pseudo Player

Suppose in a given data set, player a always wins. In this case, the denominator of equation (8) is zero, and the corresponding estimate $\hat{\omega}(a) = \infty$. An analogous case happens in the case where player a always loses, with the resulting estimate $\hat{\omega}(a) = 0$. Numerically, both cases are awkward, and some regularization is clearly warranted.

Our regularization scheme is motivated by analogy to the conjugate priors of the binomial. If our prior is a beta distribution with shape parameters α and β, then the Bayes estimate of the underlying proportion π after observing k counts of n is

$$\hat{\pi} = \frac{k + \alpha}{n + \alpha + \beta}, \tag{10}$$

or, alternately, its associated odds is $\hat{\pi}/(1 - \hat{\pi}) = (k + \alpha)/(n - k + \beta)$. This can be seen as taking two complementary counts, k and $n - k$ and adding α and β to them, respectively. Symmetry arguments suggest α and β be chosen equal, and the denominator of expression (10) establishes $\alpha+\beta$ to equal an increment in sample size. These considerations motivate the following modification of estimating equation (8):

$$1 = \frac{\sum_t y_t \, p(b_t|a,b_t) + \lambda p(a_0|a,a_0)}{\sum_t z_t \, p(a_t|a,b_t) + \lambda p(a|a,a_0)}$$

$$= \frac{\sum_t y_t \, \omega(b_t)/[\omega(a_t)+\omega(b_t)] + \lambda/[\omega(a)+1]}{\sum_t z_t \, \omega(a_t)/[\omega(a_t)+\omega(b_t)] + \lambda\omega(a)/[\omega(a)+1]}, \tag{11}$$

with the natural convention that the mojo of the pseudo-player a_0 has mojo $\omega(a_0)$ pinned at 1. Equation (11) essentially creates 2λ pseudo games with each other player a, which are won and lost in equal proportion. 2λ is typically≈ 1.

2.5 The BT Standard Error

Finally, we note that the estimates of the Bradley-Terry model are amenable to closed-form expressions of its standard errors. This calculation utilizes the delta method, and its derivation flows most easily from equation (8).

$$Var\{\log(\hat{\omega}(a))\} \doteq 1/[\sum_{t\in a} p(a_t|a_t,b_t)p(b_t|a_t,b_t)]$$

$$= 1/\sum_{t\in a} \frac{\omega(a_t)\omega(b_t)}{[\omega(a_t)+\omega(b_t)]^2}. \tag{12}$$

The interpretation of equation (12) is straightforward. The relative precision of the estimates $\hat{\omega}(a)$ is improved (a) by playing a in more slams, and (b) by playing a in slams in which it is more closely matched.

Simon and Yao (1999) offer a correction to expression (12) that accounts for the use of pseudo raters.

3 The Play-Next Rule

3.1 Objective Function

Our core idea is inspired by Dunnett's (1960) Bayes risk. Consider the linear function

$$\sum_a c(a)\ \log(\hat{\omega}_{(a)}), \tag{13}$$

where $\hat{\omega}_{(a)}$ denotes the ath smallest estimate of BT mojos $\omega_1, \omega_2, ..., \omega_n$. Taking $0 < c(1) \leq c(2) \leq ... \leq c(n)$, expression (13) allows us to emphasize the items that are ranked higher. The coefficients c also provide the means to encourage more data acquisition for the higher ranked items; we refer to the function $c(\bullet)$ the *mojovian*. Consider the variance of expression (13):

$$Var\{\sum_a c(a)\ \log(\hat{\omega}_{(a)})\}$$

$$= \sum_a c(a)^2 Var\{\log(\hat{\omega}_{(a)})\} + \sum_{a \neq b} c(a)c(b)Cov\{\log(\hat{\omega}_{(a)}), \log(\hat{\omega}_{(b)})\}. \tag{14}$$

Motivated by (12), we define the precision $v(a)$:

$$v(a) \equiv 1/Var\{\log(\hat{\omega}(a))\} = \sum_{t \in a} \frac{\omega(a_t)\omega(b_t)}{[\omega(a_t) + \omega(b_t)]^2}. \tag{15}$$

For the moment, we drop the covariance term of expression (14), it can be re-expressed as

$$\sum_a c(a)^2/v(a), \tag{16}$$

where we have implicitly aligned the indices of the functions c and v.

3.2 Most Informative Single Slam

Consider adding that single slam that pairs the performances (a, b). As a function of true parameters, the objective function (16) changes from its current value to one that benefits from this slam. In an asymptotic sense, we can assume also that the rank weights c do not change (much), and the most material difference is in the addition of the term $p(a|ab)p(b|ab) = \omega(a)\omega(b)/[\omega(a)+\omega(b)]^2$ that is added to the corresponding terms for $v(a)$ and $v(b)$. In particular,

$$\begin{aligned} & OBJFN\{now\} - OBJFN\{now + (a,b)\} \\ & = c(a)^2[\frac{1}{v(a)} - \frac{1}{v(a) + p(a|ab)p(b|ab)}] \\ & + c(b)^2[\frac{1}{v(b)} - \frac{1}{v(b) + p(a|ab)p(b|ab)}], \end{aligned} \tag{17}$$

which simplifies to

$$pq[\frac{c(a)^2}{v(a)[v(a) + pq]} + \frac{c(b)^2}{v(b)[v(b) + pq]}], \tag{18}$$

where $pq = p(a|ab)p(b|ab)$. An asymptotic argument would have $v(a) \gg pq$, so one can approximate expression (18) by

$$p(a|ab)p(b|ab)[\frac{c(a)^2}{v(a)^2} + \frac{c(b)^2}{v(b)^2}]. \tag{19}$$

A more lengthy derivation, one that accounts better for the covariance terms in objective function (14), results in this rather similar form:

$$OBJFN\{now\} - OBJFN\{now + (a,b)\} \approx p(a|ab)p(b|ab)[\frac{c(a)}{v(a)} + \frac{c(b)}{v(b)}]^2. \tag{20}$$

By maximizing the right hand side of (20), we are (modulo any approximations) identifying the single next game pair (a, b) that most reduces our variance-based objective function. To that end, we refer to

$$\{(a,b) : OBJFN\{now\} - OBJFN\{now + (a,b)\}$$
$$= \max_{(a,b)} p(a|ab)p(b|ab)[\frac{c(a)}{v(a)} + \frac{c(b)}{v(b)}]^2 \tag{21}$$

as the play-next rule.

3.3 Components of the Play-Next Rule

We propose minimizing expression (20) as a serviceable and rational approximation. The terms of (20) offer these natural interpretations:

- $pq = p(a|ab)p(b|ab) = \omega(a)\omega(b)/[\omega(a)+\omega(b)]^2$ is largest when $\omega(a) \approx \omega(b)$. This term points to the benefit of pairing players that are rather equally matched. (The opposite position is obviously false: one does not learn much about the strength of a major league baseball team from playing it against one from the farm system.)
- Larger $(c(a), c(b))$ correspond to larger $(\hat{\omega}(a), \hat{\omega}(b))$. By this property, rule (20) injects the benefit of playing higher ranked teams.
- Likewise, smaller $(v(a), v(b))$ correspond to larger standard errors of $(\hat{\omega}(a), \hat{\omega}(b))$. This element of (20) points to the benefit of playing teams for which the certainty about their mojo estimates is less.

3.4 Nomenclature for Mojovian Functions

These are all appealing properties. That said, this formulation leaves open one important issue—the choice of mojovian function. The LMPR-theory outlined in section 1.4 gives us some general guidance. The monotone nature of c, suggests consideration of

- *rank:* linearly increasing (Wilcoxon) weights, of the form $c(r) = r$;
- *reciprocal:* the reciprocal (Pareto) weights, of the form $c(r) = 1/(n-r+1)$;
- *Savage:* the so-called log-rank (Savage) weights of the form $c(1) = 1/n$; $c(r) = c(r-1) + 1/r$.

(Savage weights assign weight $1/n$ to all n performances, add weight $1/(n-1)$ to the largest $n-1$ performances, add $1/(n-2)$ to the largest $n-2$, and so on until we add $1/2$ to the top 2 performances and 1 to the highest rated one.)

The above examples of mojovian functions are all functions of the ranks of the performances; we investigate also mojovian functions that act directly on the value of ω, in particular, its linear value $c(a) = \omega(a)$ (*identity*) and square root $c(a) = \sqrt{\omega(a)}$ (*sqrt*). Rounding out the ensemble, we consider the square root of ranks (*sqrt rank*) and the constant function with $c = 1$ (*constant*).

We devote much of the remainder of this paper to exploring the best mojovian function and to comparing the scheduling using the play-next rule (20).

3.5 Aside: Reduced Memory Form

The natural algorithm to find the next-play pair would calculate (20) for all possible pairs of performances, so is of $o(n^2)$. An algorithm of smaller order is possible: It involves identifying the highest k values $c(a)/v(a)$ and recognizing that the pq-term is bounded above by $1/4$.

4 Simulations

We study the play-next rule and the value of different mojovian functions by simulation. This removes us from the direct and interesting context of the YouTube slams. At the same time, simulation-based methods allow us to investigate more algorithmic options, while minimizing the release of proprietary data.

4.1 Study Design Elements

Our simulation involves 256 videos, (1) whose $\log \omega(a)$ are drawn from the continuous logistic distribution. (b) These 256 videos are then paired into 128 initial matches, that is, each video participates in one slam with a randomly assigned opponent; the results are determined randomly in a way consistent with the Bradley-Terry probabilities. (c) Using the play-next rule, which is parametrized by different choices for the mojovian $c(\bullet)$, we form 3072 additional slam pairs, making for a total of 3200 total slam pairs. For each of steps (a), (b), and (c), we preserve the random numbers generated, thereby controlling better for the comparisons of the various mojovian functions $c(\bullet)$. We replicate steps (a), (b), and (c) five times.

The seven mojovian we assess can be divided naturally into two groups: those that are direct functions of the Bradley-Terry parameters $\omega(\bullet)$ and those that are functions of $\omega(\bullet)$ only through their ranks, $\text{rank}\, \omega(\bullet)$. Of the former category, we consider the constant function, $c(a) = 1$, the *identity* function $c(a) = \omega(a)$ and its square root $c(a) = \omega(a)^{1/2}$. For the latter, we consider the Wilcoxon (linear) *rank* function, $c(a) =\text{rank}\, \omega(a)$, its square root, $c(a) = (\text{rank}\, \omega(a))^{1/2}$, the reciprocal rank $c(a) = 1/(n + 1 - \text{rank}\, \omega(a))$, and the Savage rank, defined above. Normalized so that maxima are 1, these functions we illustrate in Figure 1.

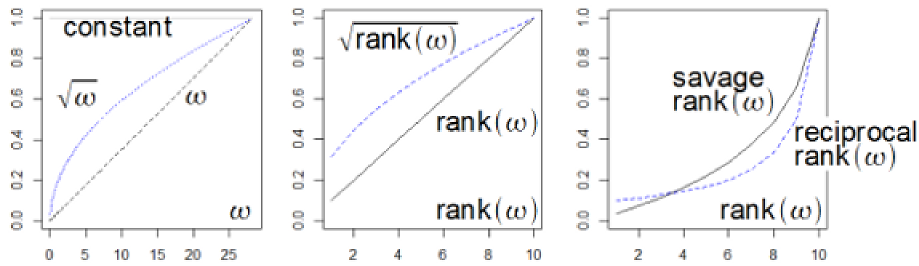

Figure 1. The seven mojovian functions investigated.

For evaluation criteria, we consider three: (1) First, as a function of estimated mojo, we want to assess the frequency with which videos are entered into plays. In this regard, we find the precision of estimated log mojo, $v(a)$ from expression (15), a useful measure. (2) Second, we would like to learn how quickly our estimates of videos converge. (3) Thirdly, from Figure 2, we observe some tendency toward bias at the high end. Therefore, we want to measure the bias associated with each choice of mojovian function.

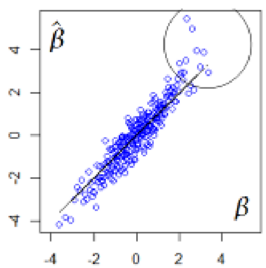

Figure 2. Estimated $\hat{\beta} \equiv \log \hat{\omega}$ plotted vs the true value $\beta \equiv \log \omega$ based on the constant mojovian. Note the bias at the higher and lower ends.

4.2 Simulation Results

Figure 3 address the first issue, how the choice of mojovian affects the slam participation over the spectrum of video strength ω. We plot log precision versus log mojo.

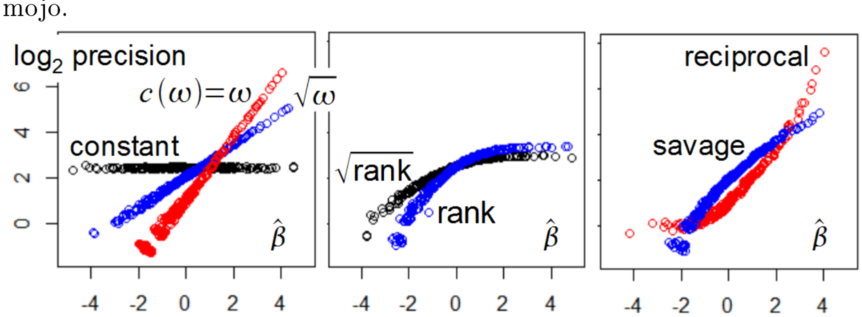

Figure 3. Precision in the form of $\log v(a)$ versus estimated $\hat{\beta} = \log \hat{\omega}(a)$. Higher values on the y-axis therefore correspond to additional slam activity.

10

As expected, the constant mojovian suggests roughly constant slam activity across the spectrum of videos and their different levels of mojo. More intriguing, the two mojovians that are powers of mojos (leftmost panel) imply a proportional relationship between slam activity and strength. In Figure 3's middle panel, the mojovian functions that are powers of the Wilcoxon ranks illustrate a different pattern: those roughly below-median mojo receive less play, while those above that level all receive about the same amount of play. By the rightmost panel of Figure 3, we see the results for the reciprocal rank and Savage score mojovian functions. Forming a convex increasing curve, the reciprocal rank score plays those at the higher end with increased activity; the convexity has the flavor of best-of-k elimination tournaments. In contrast, the curve corresponding to Savage ranks balances two properties: (1) It increases the slam views of videos with higher mojos, and (2) its concave shape suggests a sustained residual interest in the good-but-second-tier performers.

Observe also how Figure 3 charts the the slam activity of low-strength videos: Naturally, the constant $c = 1$ allocates the most attention to the low end. The mojovians of $\sqrt{\omega}$, square root rank, and (not shown) square root of reciprocal all give uncomfortably high view activity in the lower range.

In summary, Figure 3 suggests a natural ordering of the mojovian functions considered: constant, square root of rank, rank, Savage, square root of mojo, mojo (identity), and reciprocal.

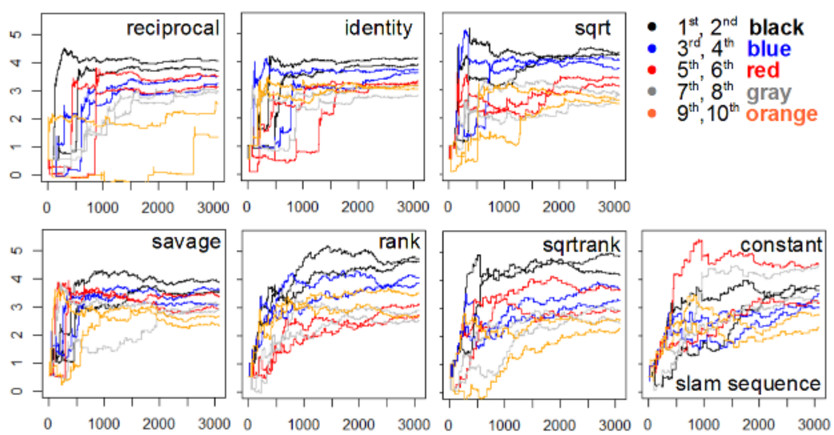

Figure 4. y-axis is estimated log mojo, $\hat{\beta} = \log \hat{\omega}$ for the 10 strongest videos; the x-axis denotes the slam performance index, ranging from 1 to 3200. Plotted here are estimates of the first of the five replicates.

In Figure 4, we see the evolving estimates of the mojos of the 10 videos with the largest mojos ω. Across various mojovians, we see generally good convergence after about 2000 slams, with notable consistency for the mojovian functions $c(a) = \omega(a)$ (identity) and Savage.

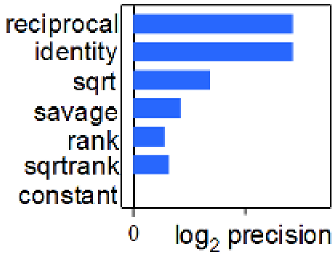

Figure 5. Relative to the constant function, the relative log precision of the estimated log mojos, for slam performance indices ≥ 2000. *Higher values are better.*

We can quantify the convergence behavior observed in Figure 4 by considering the precision of estimates with slam performance indices ≥ 2000; this summary allows us also to pool across the five simulation replicates. The result is presented in Figure 5. We observe essentially the same ordering: constant, square root of rank and rank, Savage, square root of mojo, mojo (identity), and reciprocal rank.

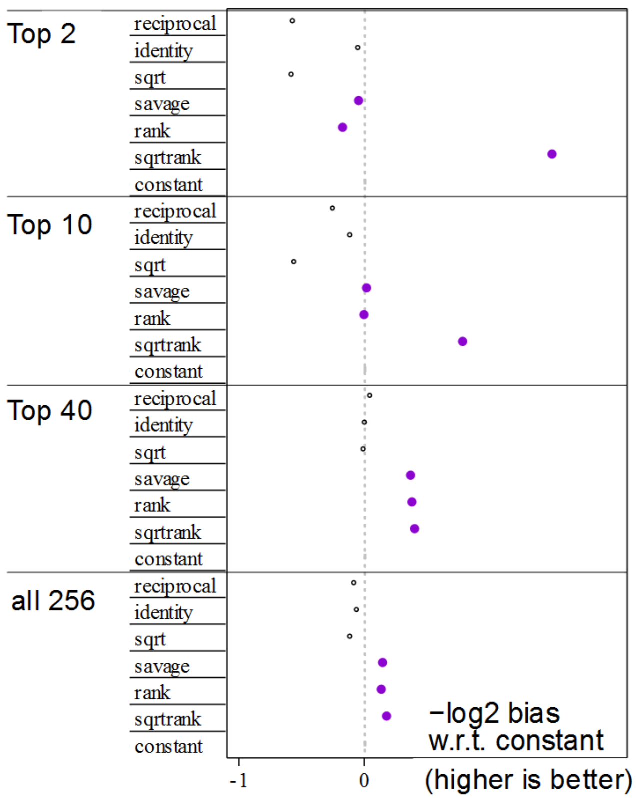

Figure 6. Relative to the constant function, the relative -log bias the 7 mojovian functions, for the top 2, *top* 10, *top* 40, *and all* 256. *Higher values are better.*

Finally, as motivated by Figure 2, one can consider the bias that the choice of mojovian induces. Figure 6 plots estimates of bias for the top 2, top 10, top 40, and finally for all 256 videos. As one can see, square root of rank shows especially good bias characteristics for the top 2 and top 10 videos. Focusing on the top 40, two almost-peers emerge, Wilcoxon rank and Savage scores.

4.3 Conclusions

This simulation study indicates that the seven mojovians considered define a spectrum. At one end, emphasizing the goal of estimating all strengths well, and, implicitly, maximizing variety, is the constant function. At the other end, emphasizing the goal of determining the best few and playing them is the reciprocal rank function. The implied order is as follows: constant, square root of rank, rank, Savage, square root of mojo, mojo (identity), and reciprocal rank.

If we have a favorite, it is the Savage (aka log-rank) score—not because it excels at any criterion, but because it strikes a comfortable median: From Figure 3 we see that it appropriately shies from showing low-mojo videos, and also that it plays the more highly ranked videos more. At the same time, it shows enough downward convexity that it implicitly encourages some level of variety. From Figures 4 and 5, we see the benefit of playing better videos more in its favorable convergence characteristics. Finally, from Figure 6, we observe that Savage scores have reasonable performance with respect to top-40 bias.

Of course, these latter remarks are intended as generic. Particular applications may require either more variety or greater high-mojo play, and therefore another choice of mojovian may be warranted.

5 An Extension

We conclude by sketching an extension to the play-next rule to non-paired conditions. For concreteness, suppose we have treatments *a, a', ...,* and we wish to identify the best few of these treatments. Further, suppose the parameters of interest is the Bernoulli probability of success, $\pi(a)$, .for each a. Designate the associated odds parameter, $\psi(a) \equiv \pi(a)/(1-\pi(a))$, its mojo.

The analogous objective function to (13) is

$$\sum_a c(a)\ \log \hat{\psi}(a), \tag{22}$$

for which we wish to minimize the variance. The delta-method approximation to the variance of the estimated log odds $\log\hat{\psi}(a)$ is $[n(a)\pi(a)(1-\pi(a))]^{-1}$. For a concrete but interesting example, take $c(a) = B(a)\pi(a)$, a proportion-domain analog to the identity function described above, times $B(a)$, the quantified benefit given the Bernoulli success. (Under this formulation, there is no monotonicity assumption, and the benefit payoff belongs to treatment a.) In this case, the variance of expression (22) becomes

$$\sum_a B(a)^2\pi(a)^2 \frac{1}{n(a)\pi(a)[1-\pi(a)]} = \sum_a \frac{B(a)^2}{n(a)} \frac{\pi(a)}{1-\pi(a)} = \sum_a \frac{B(a)^2}{n(a)}\psi(a). \tag{23}$$

By the same logic that leads to the play-next rule (19), we seek to play that treatment a such that $B(a)^2\psi(a)/n(a)^2$ is largest. In equilibrium across treatments, this implies that any treatment a would be played roughly $n(a)$ times, such that

$$n(a) \propto B(a)\sqrt{\psi(a)}. \tag{24}$$

Rule (24) is plausible, corresponding playing a treatment in proportion to its benefit and to the (square root of) estimated mojo; in particular, $n(a)/n(a') \approx B(a)/B(a')\sqrt{\psi(a)/\psi(a')}$. Note, however, that next-play rule (24) differs somewhat from the so-called probability matching rule, which suggests $n(a) \propto \pi(a)$ instead. Ignoring any $B(a)/B(a')$ term, and concentrating on $\pi(a)$-values that correspond to relatively rare events, rule (24) gives relatively flatter allocations, roughly $n(a)/n(a') \approx \sqrt{\pi(a)/\pi(a')}$.

References

[1] Blackwell, D and Girshick, MA (1954), *Theory of Games and Statistical Decisions,* Wiley, New York.

[2] Bradley, RA and Terry, ME (1952), Rank analysis of incomplete block designs I: the method of compared comparisons. *Biometrika* 39: pp 324-345.

[3] Dunnett, CW (1960), On selecting the largest of k normal population means, *Journal of the Royal Statistical Society, Series B,* 22: pp 1-40.

[4] Elo, A (1978), *The Rating of Chessplayers, Past and Present,* Arco.

[5] Gibbons, JD, Olkin, I, and Sobel, M (1999), *Selecting and ordering populations: a new statistical methodology,* Wiley, New York.

[6] Glickman, ME (1993), Paired comparison models with time-varying parameters, thesis, Harvard University, Department of Statistics.

[7] Glickman, ME (1999), Parameter estimation in large dynamic paired comparison experiments, *Applied Statistics,* 48: pp 377-394.

[8] Glickman, ME and Jensen, ST (2005), Adaptive paired comparison design, *Journal of Statistical Planning and Inference,* 127: pp 279-293.

[20] Hunter, DR (2004), MM algorithms for generalized Bradley-Terry models, *Annals of Statistics,* 32: pp 384-406.

[9] Lehmann, EL (1959), *Testing Statistical Hypotheses,* Wiley, New York.

[10] Mantel, N and Haenszel, W (1959), Statistical aspects of the analysis of data from the retrospective analysis of disease. *Journal of the National Cancer Institute,* 22: pp 719-748.

[11] Robbins, H (1952), Some aspects of the sequential design of experiments, *Bulletin of the American Mathematical Society,* 58: pp 527-535.

[12] Savage, IR (1956), Contributions to the theory of rank order statistics—the two-sample case. *Annals of Mathematical Statistics,* 28: pp 968-977.

[13] Simons, G and Yao, YC, Asymptotics when the number of parameters tends to infinity in the Bradley-Terry model for paired comparisons, *Annals of Statistics,* 27: pp 1041-1060.

[14] Thurstone, LL (1927), A law of comparative judgment, *Psychological review.*

[15] Turner, H and Firth, D (2012), Bradley-Terry models in R: the BradleyTerry2 Package, http://bradleyterry2.r-forge.r-project.org.

[16] Wald, A (1939), Contributions to the theory of statistical estimation and testing hypotheses, *Annals of Mathematical Statistics,* 11: 299-326.

[17] Wald, A (1950), *Statistical Decision Functions,* Wiley, New York.

[18] Wetherill, GB and Ofosu, JB (1974), Selection of the best of k normal populations, *Applied Statistics*, 23: 253-277.