

Bayesian Analysis for Probability-Matching

Davis Wertheimer

Abstract

Probability-matching is a phenomenon where people, given a repeated choice between alternatives with different payoff rates, select each alternative proportional to its payoff rate. We attempt to isolate the strategy responsible for this generally irrational behavior using Bayesian data analysis to score three models over three conditions. While two strategies were broadly utilized, one causing significant differences in subjects' performances over varied conditions, and the other causing few changes at all, the corresponding models did not predict subjects' responses more accurately than their competitors, indicating that we have uncovered two persistent behavioral trends in probability-matching tasks, but not its general cause.

Introduction

Although standard economic theory describes humans as rational utility-maximizers, behavioral science shows that this is often not the case. Humans make many decisions on the basis of unconscious, fast, cheap, but imprecise heuristics which in many cases counter rational behavior and logical decision-making. Some of these heuristics include over-reliance on pattern recognition, stemming from an inability to accurately perceive randomness, anchoring-and-(under)adjusting on salient cues, representative thinking and conjunction fallacies, and subadditivity (Hastie, Reid & Dawes, 2010). Many of the behaviors resulting from these heuristics are common to non-humans as well as humans, and an understanding of their function and underlying mechanisms is crucial to our understanding of the development of rational decision-making.

Probability-matching is one such heuristic, a strategy where people repeatedly pick options from a set of alternatives proportional to each options learned payoff probability. For random payoffs, this is a suboptimal strategy, because maximum expected payoff is achieved by always selecting the higher-probability option. This behavior is asymptotic, persisting over hundreds of trials (Norman & Yellott, 1966). It is also not limited to humans, and has been demonstrated in both rats and pigeons using food rewards (Robbins & Warner, 1973; Hinson & Staddon, 1983). The difference with humans is that the effect can be eliminated via a wide variety of methods, including giving the payoff probability of each option, stating explicitly that the correct answer is randomly generated, or asking subjects to give a strategy halfway through (Fantino & Esfandiari, 2002; Vulkan, 2000). This suggests that probability matching is not itself a heuristic for bandit tasks, but the direct result of other, more widely applicable, strategies or heuristics.

Two main underlying strategies have been proposed: people either attempt to exploit perceived patterns where there are none, maintaining accurate frequency representations as they do so, or they employ a noisy anchor-adjust heuristic where a correct answer reinforces the current choice and

an incorrect answer prompts switching to the other option (Fantino & Esfandiari, 2002). We propose a third strategy, where people track payoff rates for simple strategies such as stay or switch, and probability match to these strategy payoffs. This allows for more flexible behavior, especially in the case where the boxes pay off evenly but predetermined patterns dominate the payoff distribution. All three of these strategies (anchor-adjust, frequency-matching, and strategy-matching) can be represented as generative models in a Bayesian data analysis scheme.

We were also interested in whether people perform posterior reasoning about the correct answer sequence as they progress through repeated trials. We hypothesize that strategy usage varies depending on the level of perceived randomness in the answer sequence. We used Falk and Konold's subjective randomness metric, demonstrated in their 1997 paper to be a good objective predictor of subjective randomness judgments, to measure perceived randomness in our answer sequences. This metric proceeds through a binary sequence and adds one to the current score for every sequence of repeated values, and two for every sequence of alternating values. We used this metric to inform our three strategy models during our analysis.

We use these models and Bayesian data analysis to answer three questions about probability-matching behavior:

- 1) What noise level is optimal for a deterministic anchor-adjust strategy model?
- 2) Which of the three strategies best models human performance on probability-matching tasks?
- 3) To what degree do humans adjust their strategy usage based on perceived randomness in the sequence of correct answers?

Methods

Our probability-matching task, housed in a web module, consists of a choice between two labeled boxes, one of which holds an obscured marble.¹ Subjects are asked to guess which box holds the marble over 100 sequential trials with 1-second delays, in which they are told the outcome of the previous trial and shown which box actually held the marble. The trials were followed by an optional short-response questionnaire about strategy usage and preceded by a short tutorial which explained the game and asked the subject to make a series of example selections.

We split the task over three conditions, each designed to elicit different behaviors under different strategies. The "normal" condition consisted of the traditional probability-matching setup, with the right-hand box holding the marble

¹Code for the web module is available at <https://github.com/daviswert/Bayesian-Analysis-Prob-Matching/tree/master/Exp>

80% of the time, and the left-hand box holding it for the other 20%. This was the control condition, designed to elicit normal probability-matching behavior.

The "twitchy" condition consisted of the same setup, but with both box payoffs approximately equal. Instead, the answer for 80% of the time was whichever box didn't hold the marble during the previous trial. This led the marble to often switch back and forth from one box to the other. The twitchy condition is designed to confound the anchor-adjust strategy, which performs below chance in this condition, and to a lesser degree the frequency-matching strategy, which performs at chance. Strategy-matching is expected to perform at the level of normal probability-matching.

The "streaky" condition is identical to the twitchy condition except that the probabilities are reversed: the ball stays wherever it last was 80% of the time, and switches for the other 20. The streaky condition is built to confound the frequency-matching strategy, but encourage the anchor-adjusting and strategy-matching, which both perform at matching level. The streaky condition is also designed to appear less random, in order to contrast any randomness-conditioned behaviors against those appearing in the twitchy condition, which, by containing many alternations, appears more random.

We presented the task to 27 Stanford undergraduates, 9 male, 16 female, 2 declined response. 7 were in the normal condition, 10 were in the twitchy condition, and 10 were in the streaky condition.

Analysis and Results

We measured subjects' performance using two metrics: how often they were correct in their predictions, and how often they employed the strategy that yields optimal payoff. Standard probability-matching behavior entails 68% prediction accuracy and 80% optimal strategy utilization. Normal and streaky conditions showed slightly below probability-matching performance, with a significant drop-off in the twitchy condition for both prediction accuracy (1-way ANOVA, $p=.045$) and maximization (1-way ANOVA, $p=.005$). Results given in table 1 and fig. 1.

Table 1: Performance on Probability-Matching Task

Condition	Prediction Accuracy	Maximization
Normal	65.7%	76.4%
Twitchy	63.2%	64.8%
Streaky	70.2%	73.6%

For the Bayesian analysis, we built enumeration queries for each strategy model that returned the posterior distribution of the strategy accurately predicting a subject's response for any given condition, trial, and subject.² These

²Code for the generative models and subsequent analyses is available at <https://github.com/daviswert/Bayesian-Analysis-Prob-Matching/blob/master/Final%20analysis%20code.txt>

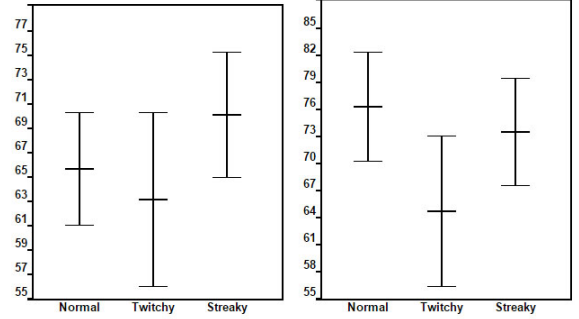


Figure 1: Subject performance across conditions. Left-hand chart gives prediction accuracy, right-hand chart gives maximization frequency.

queries are parametrized by aggregate information about the subject's response history and the answer sequence, so each query returns a conditional correctness probability $P(X_n|X_1, X_2, \dots, X_{n-1})$ where X_n represents the choice made on trial n . Ideally, what we would like to do is condition on a strategy model accurately predicting a subject's entire response history, which equates to a product of the individual conditional queries:

$$P(X_1, X_2, \dots, X_{100}) = P(X_1)P(X_2|X_1)P(X_3|X_2, X_1) \dots P(X_{100}|X_1, X_2, \dots, X_{99})$$

Unfortunately, the resulting probability is so low that not only can we not feasibly condition on it, it is impossible to calculate due to underflow errors. We then tried to add the log-probabilities instead, in order to reach an equivalent formulation, but the log-probabilities generated large enough negative numbers to cause overflow when summed over 100 trials. Finally, we used factor statements on each strategy's average score, represented as probability of correct prediction, over all 100 trials. We thus score our models using the following formula:

$$S = \frac{1}{100} \sum_{i=1}^{100} P(X_i|X_1, X_2, \dots, X_{i-1})$$

This represents a strategy model's average ability to predict any particular choice in a subject's response history, given complete knowledge of performance up to that point.

Throughout our analysis we use only enumeration queries and factor statements, often by necessity, but because of this our analyses involve no random sampling. We calculate the real posterior distributions on every query for the data provided.

Question 1

What noise level is optimal for a deterministic anchor-adjust strategy model? We calculated separate answers to this question for each condition, and then aggregate across conditions. For each subject in the condition, we used enumeration query to sample noise values as percentages ranging from 0.0 to 0.9. We then passed this parameter through the scoring function for the anchor-adjust strategy model directly

to the model itself, which for any given choice randomly probability-matched with probability given by the noise percentage, and applied deterministic anchor-adjust calculations otherwise. The outer enumeration query factored noise levels based on the score of the strategy model set to that noise level, and posterior noise distributions were summed over all subjects in the condition, since subjects' performances are independent. Results are given in fig. 2.

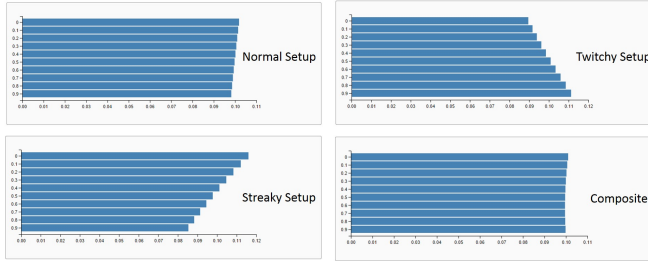


Figure 2: Noise level distributions for optimal performance of the anchor-adjust model across conditions.

Results are characteristic of strategy maximization. In cases where anchor-adjust outperforms frequency-matching, optimal noise level is set to 0. Otherwise, frequency-matching is optimal, and the optimal noise level is 1. Without a selection criterion, our question amounts to deciding which strategy is globally optimal, the answer to which varies by condition. We address this question of selection criteria in Question 3. For the purposes of subsequent analysis, we assume a noise level of 0, in order to keep the anchor-adjust strategy entirely distinct from frequency-matching.

Question 2

Which of the three strategies best models human performance on probability-matching tasks? We again calculated a separate answer to this question for each condition, and then on the aggregate data. For each subject in the condition, we used enumeration query to sample evenly from each of the three strategy models. We then calculated the score for that model on the given subject's response sequence, and factored the modeled strategy based on its score. Posterior strategy distributions were summed over all subjects in the condition, since subjects' performances are independent. Results are given in fig. 3.

Results vary widely by condition, indicating that subjects are extremely flexible and adaptive. Optimal strategies are employed most often for each condition, with anchor-adjust leading in the streaky condition and tying with frequency-matching in the normal condition, and strategy-matching leading in the twitchy condition. Strategy-matching is used very close to or slightly more than chance (one third of the time) within every condition, making it the most robust of the three strategy models. Anchor-adjusting varies widely according to its degree of optimality in the condition, making it the most flexibly implemented strategy model of the three.



Figure 3: Performance comparisons of the three strategy models across conditions. 1 corresponds to anchor-adjust, 2 corresponds to frequency-matching, and 3 corresponds to strategy-matching.

Question 3

To what degree do humans adjust their strategy usage based on perceived randomness in the sequence of correct answers? To answer this question we aggregated data across conditions, since the Falk and Konold metric remained highly consistent between subjects within conditions (figure not included due to lack of visually communicable variance). Because F&K increases linearly with the length of the sequence, we divide trial n 's complexity score by n in order to obtain a complexity score that can be compared across trials. Average final complexity scores for the normal, twitchy, and streaky conditions are .296, .397, and .206 respectively.

In order to judge complexity effects, we implemented a scoring model that takes two strategies, and for every trial generates both of their posterior probability distributions of a correct answer. It then factors each strategy's posterior according to its probability of being selected, given by the current complexity score, and takes the average over all trials as per the scoring function. This produces the score for a model which selects one strategy in perceived highly random contexts, and the other in non-random ones.

This model predicts greater utilization of the high-randomness-assigned strategy in early trials, before the randomness score can converge to its final asymptotic value, and a longer burn-in for high-perceived-randomness conditions like twitchy.

We compared this "complex" model against a "noisy" model, which samples from the same two strategies but uses the average complexity over all trials instead of the perceived accuracy at that moment in time. This model yields the same aggregate distribution of strategy selection over all trials as its complex counterpart, but decouples the selection mechanism from the perceived randomness of the answer sequence. We then sampled between these two metamodels using enumeration query, and factored on their respective scores. The resulting distribution shows to what degree the complex model is more accurate than the noisy one, and to what degree subjects alter strategy usage based on perceived randomness. Results are given in fig. 4.

For each strategy model under consideration, we tested the complex metamodel against the noisy metamodel using that strategy as the high-randomness-assigned strategy and totally

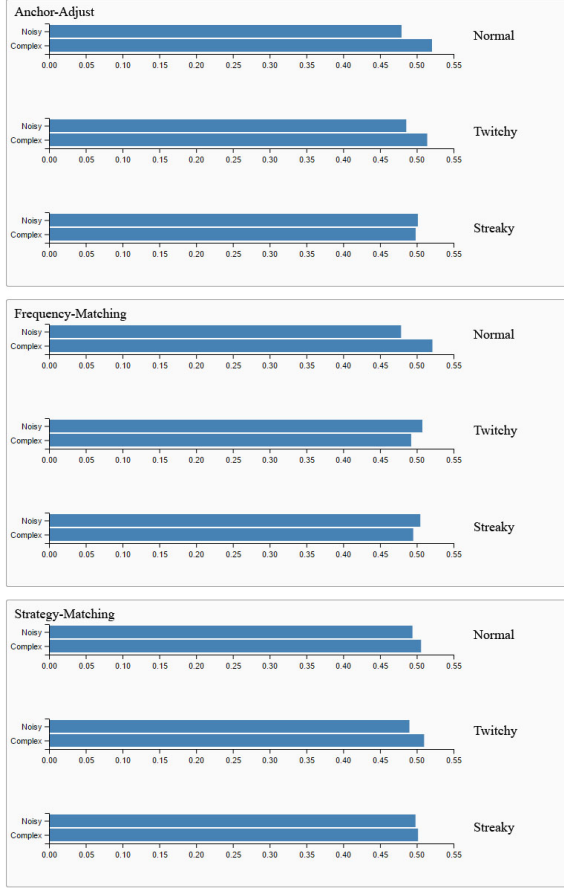


Figure 4: Comparisons between complex metamodel and noisy metamodel over all strategies and conditions. Both models employ the given strategy for high randomness trials and random sampling for low randomness trials. The degree to which the complex model outperforms the noisy model is the degree to which the given strategy is relied upon during trials of perceived high randomness.

random sampling as the low-randomness-assigned strategy. Switching the order simply reverses the bars. Scaling the complexity probability by a constant scales the differences already present by the same constant, and adding a constant value to the probability has no effect. In order to evaluate a complex model that employs one strategy model and then another (not random sampling), simply add the distribution of the first strategy to the inverse distribution of the second.

The one persistent trend visible in the data is that the complex metamodel using strategy-matching consistently outperforms the noisy equivalent. This is only by a small margin, but this margin can be arbitrarily scaled up to three times higher by adjusting the function that converts the complexity score into a sampling probability.

Discussion

While the results are largely ambiguous, we are able to extract some valuable pieces of information. First, examining the non-Bayesian data analysis, there was a significant drop in both performance and maximization for the twitchy condition, and a slightly smaller one for the normal condition versus the streaky one. This gradual performance drop coincides both with an increase in perceived randomness, and a decrease in the utilization of the anchor-adjust strategy. While it is tempting to correlate anchor-adjust utilization with perceived randomness, we must keep in mind that the optimality of the anchor-adjust strategy also decreases across these conditions. It is far more likely that subjects are simply learning not to use the anchor-adjust strategy. However, the distribution of strategy models for the twitchy condition shows the anchor-adjust model as performing much more highly than would be expected, given that it predicts below-chance performance. This, combined with the significant performance drop observed in the twitchy condition, indicates that an anchor-adjust bias does exist, and does contribute to probability-matching behavior, but that it is not a powerful predictor of fine-grained human behavior in probability-matching tasks.

Second, the strategy-matching model is a consistently average predictor of human behavior on our probability-matching task. Thus, while it is not a particularly powerful predictor, it is highly robust, and accommodates the flexible learning behavior observed between the three conditions. It is also the only strategy whose usage is consistently conditioned on the perceived randomness of the answer sequence. This indicates that there is a Bayesian learning process occurring, but that we have not fully managed to capture it with the simple strategies presented here.

While powerful, the Bayesian data analysis scheme itself presents some problems for our results. Because we are using factor statements instead of hard conditions, we observe a regression toward the mean effect, which compounds recursively for each metamodel we implement. That we use sums of individual choice distributions in our generative models, rather than their products, contributes to this regression effect, because a setup that uses products to generate the distribution for the entire 100-trial sequence would allow errors in the model to compound as they should. As is, they do not. This makes it difficult to determine which effects are significant, and which are not. On the one hand, the data presents us with very small differences in strategy model performances. On the other hand, we can assume a high degree of regression toward the mean, and must keep in mind that the data we are working with does not represent a random, possibly skewed sample of model performance, but the true posterior.

Conclusion

Using Bayesian data analysis, we have isolated two weak behavioral trends in human performance on probability-matching tasks. First, we observe general anchor-adjust bi-

ases, which contribute to probability-matching behavior, and decrease performance in situations where such a strategy is not optimal. Second, we can demonstrate that subjects are using Bayesian inference to modulate their strategy learning. This strategy learning process is reflected in our data by the fact that strategy scores vary widely between conditions, except for strategy-matching, which, as we have implemented it, is a very rudimentary form of Bayesian inference on strategy usage. We suspect that we were unable to fully model learning behavior because the learned strategies are much more complex than the simple stay versus switch model presented here. Attempting to implement more robust models of strategy learning would be a profitable direction for future work. Integrating our findings on strategy inference with the resource-rational model presented by Vul, Goodman, Griffiths, and Tenenbaum in their 2014 paper could also prove profitable, especially when it comes to using n-back sampling as opposed to the full aggregate answer history to inform individual trial decisions and predictions for a single subject.

Acknowledgements

Special thanks to Noah Goodman and Michael Henry Tessler for inspiration and plenty of help on this project.

References

- Edward Vul, T. L. G., Noah D. Goodman, & Tenenbaum, J. B. (2014). One and done? optimal decisions from very few samples. *Cognitive Science*, 4, 599–637.
- Falk, R., & Konold, C. (1997). Making sense of randomness: Implicit encoding as a basis for judgment. *Psychological Review*, 104.
- Fantino, E., & Esfandiari, A. (2002). Probability matching: Encouraging optimal responding in humans. *Canadian Journal of Experimental Psychology*, 56.
- Hastie, R., & Dawes, R. (2010). *Rational choice in an uncertain world: The psychology of judgment and decision making*. London: SAGE.
- Hinson, J. M., & Staddon, J. E. R. (1983). Matching, maximizing, and hill-climbing. *Journal of the Experimental Analysis of Behavior*, 40.
- Norman, M. F., & Yellott, J. I. (1966). Probability matching. *Psychometrika*, 31.
- Robbins, D., & Warner, P. L. (1973). Individual organism probability matching with rats in a two-choice task. *Bulletin of the Psychonomic Society*, 2.
- Vulkan, N. (2000). An economist's perspective on probability matching. *Journal of Economic Surveys*, 14.