Bounding Consideration Probabilities in Consider-Then-Choose Ranking Models

Extended Abstract

Ben Aoki-Sherwood
Johns Hopkins Applied Physics Lab
Laurel, MD, United States
aokisherwoodb@gmail.com

Catherine Bregou
Carleton College
Northfield, MN, United States
bregouc@carleton.edu

David Liben-Nowell
Carleton College
Northfield, MN, United States
dln@carleton.edu

Kiran Tomlinson
Cornell University
Ithaca, NY, United States
kt@cs.cornell.edu

Thomas Zeng
University of Wisconsin
Madison, WI, United States
tpzeng@cs.wisc.edu

ABSTRACT

A common theory of choice posits that individuals make choices in two steps, first selecting a subset of the alternatives to consider before making a choice from the resulting consideration set. However, inferring unobserved consideration sets (or item consideration probabilities) in this "consider then choose" setting poses significant challenges: even simple models of consideration with strong independence assumptions are not identifiable, even if item utilities are known. We consider a natural extension of consider-then-choose models to a top-k ranking setting, where we assume rankings are constructed according to a Plackett–Luce model after sampling a consideration set. While item consideration probabilities remain non-identified in this setting, we prove that knowledge of item utilities allows us to infer bounds on the relative sizes of consideration probabilities. Additionally, given a bound on the expected consideration set size, we derive absolute upper and lower bounds on item consideration probabilities. We also provide an algorithm to tighten those bounds on consideration probabilities by propagating inferred constraints. Thus, we show that we can learn useful information about consideration probabilities despite their non-identifiability. We demonstrate our methods on a dataset from a psychology experiment with two different ranking tasks (one with fixed consideration sets and one with unknown consideration sets). This combination of data allows us to estimate utilities and then learn about unknown consideration probabilities using our bounds.

CCS CONCEPTS

• **Applied computing → Law, social and behavioral sciences**; • **Theory of computation → Theory and algorithms for application domains**; • **Information systems → Learning to rank**.

KEYWORDS

discrete choice models; Plackett–Luce; consideration sets

Proc. of the 23rd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2024), N. Alechina, V. Dignum, M. Dastani, J.S. Sichman (eds.), May 6 – 10, 2024, Auckland, New Zealand. © 2024 International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org).

ACM Reference Format:
Ben Aoki-Sherwood, Catherine Bregou, David Liben-Nowell, Kiran Tomlinson, and Thomas Zeng. 2024. Bounding Consideration Probabilities in Consider-Then-Choose Ranking Models: Extended Abstract. In *Proc. of the 23rd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2024), Auckland, New Zealand, May 6 – 10, 2024*, IFAAMAS, 3 pages.

Background. Among the wide-ranging topics studied in the behavioral sciences, predicting and explaining human decisions is a central challenge. (Why *that* entrée at *that* café with *that* person, of all the restaurants you can think of, of all your potential dates, of all the dishes on the menu?) Settings where individuals select an item from a collection of available alternatives are well-studied in the literature on *discrete choice* [30], with applications ranging from marketing [8] and voting [29] to transportation policy [13] and recommender systems [9]. A closely related line of work studies *rankings* [2], often modeling a ranking as a sequence of discrete choices (selecting the top-ranked item, then the second, etc.). The *Plackett–Luce ranking model* [15, 22] exemplifies the tight link between discrete choice and ranking, positing that items at each of k positions are selected in turn according to a logit choice model [18].

In keeping with the framework of bounded rationality [27], a prominent line of work in discrete choice suggests that selection is a two-stage process, where individuals first narrow their options to a small *consideration set* from which a final selection is made [12, 26]. These "consider then choose" models have shown considerable promise [4, 24], but suffer from the unobservability of consideration sets (except in carefully controlled experimental settings) and cannot in general be identified from observed choice data [14, 31]. Perhaps as a result, consider-then-choose models have received comparatively little attention in the ranking literature.

Plackett–Luce with consideration: the PL+C model. Here, we study the natural consider-then-rank model obtained by augmenting top-k Plackett-Luce with the independent-consideration rule [17], which we term PL+C. Each of the n items in a *universe* $\mathcal{U}$ has an item-specific *consideration probability* $p_i \in (0, 1]$ and a *utility* $u_i \in \mathbb{R}$. The two-stage choice unfolds as follows, for fixed $k \leq n$:

(1) Consideration: Each item $i \in \mathcal{U}$ advances to the second stage randomly and independently with probability p_i, yielding some *consideration set* $C \subseteq \mathcal{U}$. (We do not observe cases where $|C| < k$—either such instances are thrown out before we observe them, or the chooser resamples a new set.)

(2) **Ranking:** A length-k ranking is formed by a sequence of k choices; the probability of selecting i is proportional to $\exp(u_i)$, first choosing a top-ranked item, then a second-ranked item (distinct from the first), etc. The choice in PL+C is made from among the elements of C. (In the classical Plackett–Luce model, *all* elements of $\mathcal{U}$ are considered.)

We imagine making observations of many individuals' rankings, hoping to infer consideration probabilities and/or utilities. One might hope that richer observations—a k-item ranking rather than just a single choice—would make it feasible to identify consideration probabilities, at least for large k. Unfortunately, it does not:

Theorem 1. *For all $n \geq 1$ and $1 \leq k \leq n$, there exist multiple sets of consideration probabilities that generate exactly the same distribution over rankings, with fixed utilities u_i for each $i \in \mathcal{U}$.*

However, we show that it is nonetheless possible to derive meaningful bounds. In addition to observations of rankings, we draw on two types of information for learning consideration probabilities: (1) known item utilities and (2) a lower bound on expected consideration set size. The intuition is that if i has higher utility then j, then we would expect to see i ranked highly more often than j—but if instead j outperforms i, then consideration must be the culprit.

Theorem 2. *For items $i, j \in \mathcal{U}$ with $u_i > u_j$, if, for some $\ell \geq 1$, item i appears less often in top-ℓ rankings than item j, then $p_j > p_i$. (If i appears $c < 1$ times as often as j, then $p_i/(1-p_i) \leq c \cdot p_j/(1-p_j)$.)*

Theorem 2 allows us to propagate an upper bound on p_j into one on p_i, and a lower bound on p_i into one on p_j. To get started, we also seek absolute upper and lower bounds on consideration probabilities. (It is tempting to think that item i's rate of occurrence in observed top-k rankings would lower bound p_i, but that bound does not hold, because PL+C conditions samples on $|C| \geq k$. However, with relatively mild assumptions—see (2), above—item occurrence rates still yield a lower bound on consideration probabilities.) Using broadly similarly techniques, we derive upper bounds, too. (The technical infrastructure is more involved here; we derive bounds on the counterintuitive impact of the effect in PL+C that increasing p_j can actually make item $i \neq j$ *more* likely to be chosen, not less.)

Theorem 3. *Suppose $\sum_{i \in \mathcal{U}} p_i \geq \alpha k$ for some $\alpha > 1$. For any $i \in \mathcal{U}$, the consideration probability for i is bounded as*

$$p_i \geq (\textit{frequency } i \textit{ is ranked in the top } k) \cdot [1 - \varepsilon_1(\alpha, k)] \textit{ and}$$

$$p_i \leq f(u_i) \cdot [(\textit{frequency } i \textit{ is ranked in first place}) + \varepsilon_2(\alpha, k)]$$

for a function f and small functions $\varepsilon_1, \varepsilon_2$ (decreasing in α and k).

Finally, we provide efficient algorithms that tighten our absolute bounds (Theorem 3) by propagating them over a directed acyclic graph induced by our relative bounds (Theorem 2). Details of theorems and algorithms are available in the full version of the paper.

Application to empirical data. Finally, we present an illustration of how our theoretical machinery can reveal meaningful information about consideration probabilities on real data. In an existing psychology experiment about perceptions of U.S. history [23], participants completed several tasks related to their view of the historical importance of particular U.S. states, such as naming the three states they believe contributed most to U.S. history. They were also provided with a random set of 10 states and asked to rate their percentage contribution to U.S. history. (We convert the numerical scores to rankings.) These two settings, *Top-3* and *Random-10*, provide us both observations from fixed consideration sets (Random-10) and from unknown consideration sets (Top-3), in which participants must recall the names of the states that they will list [6]. We thus estimate utilities in the absence of consideration from the Random-10 data, and then estimate consideration from the Top-3 data, using the PL+C model and our bounds (implemented in Pytorch [21]).

The algorithms described above rely on the existence of pairs of states whose utilities and top-ℓ ranking rates are flipped. Interestingly, many such flips occur in this data, highlighting the apparent importance of consideration in the Top-3 question. Using our methods, we find that, e.g., $p_{\text{Massachusetts}} > p_{\text{Virginia}} > p_{\text{Pennsylvania}}$ and $p_{\text{Virginia}} > p_{\text{New York}}$. Combining our lower and upper bounds yields feasible intervals on consideration probabilities for each state, revealing that, if our assumptions are valid, most states are considered less than 30–40% of the time. Additionally, the bounds on consideration probabilities align with theories about why certain states were highly rated in the data [23].

Related work. Prior research on consideration focused on single choices rather than rankings. The approach most closely related to our work adds a consideration stage to random utility models [4, 5, 24, 31], following Manski's formulation [16], but there are many alternative strategies [5, 7, 14, 17, 24]. Existing discrete choice approaches to handle non-identifiability of consideration probabilities use explicit item availability questions [5, 24, 28], online browsing data [11, 19], observations of "none of the above" outside options [14, 17], item features that change over time [1], and parametric models of consideration [4, 31]. Consideration has received only limited attention in the ranking literature [10, 20].

Discussion. We formalized a natural model of ranking with consideration, augmenting Plackett–Luce with an independent consideration model. Despite showing that consideration probabilities are not identified in general, we derived relative and absolute bounds that allow us to learn about possible ranges of consideration probabilities from observed ranking data. Our data application demonstrates how these bounds can be used in practice to gain insight into consideration behavior from ranking data.

There remains much to explore regarding the PL+C model, and ranking with consideration sets more generally. First, a thorough characterization of PL+C would be valuable, including its expressive power relative to other augmented Plackett–Luce models, such as the contextual repeated selection model [25]. Another interesting question concerns computing PL+C probabilities efficiently. If we know utilities and consideration probabilities, the direct approach to computing ranking probabilities involves a sum over exponentially many possible consideration sets. Is it possible to compute PL+C probabilities in polynomial time, or is it provably hard?

For further discussion of related work, limitations, future directions, and technical details, we refer readers to the full paper [3].

Portions of this work were performed while all authors were at Carleton College. Thanks to Aadi Akyianu, David Chu, Katrina Li, Adam Putnam, Sophie Quinn, Morgan Ross, Laura Soter, and Jeremy Yamashiro for helpful discussions.

REFERENCES

[1] Jason Abaluck and Abi Adams-Prassl. 2021. What do consumers consider before they choose? Identification from asymmetric demand responses. *The Quarterly Journal of Economics* 136, 3 (2021), 1611–1663.

[2] Mayer Alvo and Philip LH Yu. 2014. *Statistical Methods for Ranking Data*. Springer, New York.

[3] Ben Aoki-Sherwood, Catherine Bregou, David Liben-Nowell, Kiran Tomlinson, and Thomas Zeng. 2024. Bounding Consideration Probabilities in Consider-Then-Choose Ranking Models. arXiv:2401.11016 [cs.LG] https://arxiv.org/abs/2401.11016

[4] Gözen Başar and Chandra Bhat. 2004. A parameterized consideration set model for airport choice: an application to the San Francisco Bay area. *Transportation Research Part B: Methodological* 38, 10 (2004), 889–904.

[5] Moshe Ben-Akiva and Bruno Boccara. 1995. Discrete choice models with latent choice sets. *International Journal of Research in Marketing* 12, 1 (1995), 9–24.

[6] John Brown (Ed.). 1976. *Recall and Recognition.* John Wiley & Sons, London.

[7] Matias D Cattaneo, Xinwei Ma, Yusufcan Masatlioglu, and Elchin Suleymanov. 2020. A random attention model. *Journal of Political Economy* 128, 7 (2020), 2796–2836.

[8] Pradeep K Chintagunta and Harikesh S Nair. 2011. Discrete-choice models of consumer demand in marketing. *Marketing Science* 30, 6 (2011), 977–996.

[9] Mazen Danaf, Felix Becker, Xiang Song, Bilge Atasoy, and Moshe Ben-Akiva. 2019. Online discrete choice models: Applications in personalized recommendations. *Decision Support Systems* 119 (2019), 35–45.

[10] Dennis Fok, Richard Paap, and Bram Van Dijk. 2012. A rank-ordered logit model with unobserved heterogeneity in ranking capabilities. *Journal of Applied Econometrics* 27, 5 (2012), 831–846.

[11] Bin Gu, Prabhudev Konana, and Hsuan-Wei Michelle Chen. 2012. Identifying consumer consideration set at the purchase time from aggregate purchase data in online retailing. *Decision Support Systems* 53, 3 (2012), 625–633.

[12] John R Hauser and Birger Wernerfelt. 1990. An evaluation cost model of consideration sets. *Journal of Consumer Research* 16, 4 (1990), 393–408.

[13] Matt Horne, Mark Jaccard, and Ken Tiedemann. 2005. Improving behavioral realism in hybrid energy-economy models using discrete choice studies of personal transportation decisions. *Energy Economics* 27, 1 (2005), 59–77.

[14] Srikanth Jagabathula, Dmitry Mitrofanov, and Gustavo Vulcano. 2023. Demand estimation under uncertain consideration sets. *Operations Research* (2023).

[15] R Duncan Luce. 1959. *Individual Choice Behavior: A Theoretical Analysis.* Wiley, New York.

[16] Charles F Manski. 1977. The structure of random utility models. *Theory and Decision* 8, 3 (1977), 229.

[17] Paola Manzini and Marco Mariotti. 2014. Stochastic choice and consideration sets. *Econometrica* 82, 3 (2014), 1153–1176.

[18] Daniel McFadden. 1973. Conditional logit analysis of qualitative choice behavior. In *Frontiers in Econometrics*, Paul Zarembka (Ed.). Academic Press, New York, 105–142.

[19] Wendy W Moe. 2006. An empirical two-stage choice model with varying decision rules applied to internet clickstream data. *Journal of Marketing Research* 43, 4 (2006), 680–692.

[20] Marco A Palma. 2017. Improving the prediction of ranking data. *Empirical Economics* 53 (2017), 1681–1710.

[21] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, et al. 2019. Pytorch: An imperative style, high-performance deep learning library. *Advances in Neural Information Processing Systems* 32 (2019), 8026–8037.

[22] Robin L Plackett. 1975. The analysis of permutations. *Journal of the Royal Statistical Society Series C: Applied Statistics* 24, 2 (1975), 193–202.

[23] Adam L Putnam, Morgan Q Ross, Laura K Soter, and Henry L Roediger III. 2018. Collective narcissism: Americans exaggerate the role of their home state in appraising US history. *Psychological Science* 29, 9 (2018), 1414–1422.

[24] John H Roberts and James M Lattin. 1991. Development and testing of a model of consideration set composition. *Journal of Marketing Research* 28, 4 (1991), 429–440.

[25] Arjun Seshadri, Stephen Ragain, and Johan Ugander. 2020. Learning rich rankings. *Advances in Neural Information Processing Systems* 33 (2020), 9435–9446.

[26] Allan D Shocker, Moshe Ben-Akiva, Bruno Boccara, and Prakash Nedungadi. 1991. Consideration set influences on consumer decision-making and choice: Issues, models, and suggestions. *Marketing Letters* 2 (1991), 181–197.

[27] Herbert A Simon. 1957. *Models of Man: Social and Rational.* Wiley, New York.

[28] Jung-Chae Suh. 2009. The role of consideration sets in brand choice: The moderating role of product characteristics. *Psychology & Marketing* 26, 6 (2009), 534–550.

[29] Paul W Thurner. 2000. The empirical application of the spatial theory of voting in multiparty systems with random utility models. *Electoral Studies* 19, 4 (2000), 493–517.

[30] Kenneth E Train. 2009. *Discrete Choice Methods with Simulation.* Cambridge University Press, Cambridge.

[31] Erjen van Nierop, Bart Bronnenberg, Richard Paap, Michel Wedel, and Philip Hans Franses. 2010. Retrieving Unobserved Consideration Sets from Household Panel Data. *Journal of Marketing Research* 47, 1 (2010), 63–74.