



Combinatorial Inference on the Optimal Assortment in the Multinomial Logit Model

SHUTING SHEN, Duke University, USA

XI CHEN, New York University, USA

ETHAN X. FANG, Duke University, USA

JUNWEI LU, Harvard T.H. Chan School of Public Health, USA

Assortment optimization has received active explorations in the past few decades due to its practical importance. Despite the extensive literature dealing with optimization algorithms and latent score estimation, uncertainty quantification for the optimal assortment still needs to be explored and is of great practical significance. Instead of estimating and recovering the complete optimal offer set, decision-makers may only be interested in testing whether a given property holds true for the optimal assortment, such as whether they should include several products of interest in the optimal set, or how many categories of products the optimal set should include. This paper proposes a novel inferential framework for testing such properties. We consider the widely adopted multinomial logit (MNL) model, where we assume that each customer will purchase an item j within the offer set of products $\mathcal{S}$ with a probability proportional to the underlying preference score u_j^* associated with the product. For a full assortment of n products, our objective is to conduct a hypothesis test concerning a general optimal assortment property, given by:

$$H_0 : \mathcal{S}^* \in \mathcal{S}_0 \text{ versus } H_1 : \mathcal{S}^* \notin \mathcal{S}_0,$$

where $\mathcal{S}^*$ denotes the optimal offer set, and $\mathcal{S}_0$ is a set of offer sets satisfying the property of interest. We reduce inferring a general optimal assortment property to quantifying the uncertainty associated with the sign change point detection of the marginal revenue gaps defined as $\Delta_k = \sum_{i=1}^k r_i u_i^* - (\sum_{i=0}^k u_i^*) r_k$, $k = 1, \dots, n$, where $r_1 \geq \dots \geq r_n$ are the revenue parameters associated with the n products. By plugging in the Newton-debiased maximum likelihood estimator (MLE) for the latent preference scores, we obtain the marginal revenue gap estimators $\hat{\Delta}_k$'s and show their asymptotic normality. Furthermore, we construct a maximum statistic via the gap estimators to detect the sign change point:

$$T = \max_{k \in [n]} (\hat{\sigma}_k^2 / L)^{-1/2} (\hat{\Delta}_k - \Delta_k),$$

where $\hat{\sigma}_k^2$ is a consistent estimator for the asymptotic variance of $\hat{\Delta}_k$. By approximating the distribution of the maximum statistic with multiplier bootstrap techniques, we propose a valid testing procedure. We also conduct numerical experiments to assess the performance of our method.

Xi Chen would like to thank the support from NSF via the Grant IIS-1845444. Ethan X. Fang would like to thank the support from NSF DMS-2230795 and NSF DMS-2230797. Junwei Lu would like to thank the support from NSF 1916211, NIH 1R35CA220523-01A1, and NIH 5U01CA209414-02.

A full version of this paper can be found at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4371919.

CCS Concepts: • **Computing methodologies** → **Machine learning**; • **Mathematics of computing** → **Probability and statistics**; • **Applied computing** → **Operations research**.

ACM Reference Format:

Shuting Shen, Xi Chen, Ethan X. Fang, and Junwei Lu. 2023. Combinatorial Inference on the Optimal Assortment in the Multinomial Logit Model. In *Proceedings of the 24th ACM Conference on Economics and Computation (EC '23)*, July 9–12, 2023, London, United Kingdom. ACM, New York, NY, USA, 1 page. <https://doi.org/10.1145/3580507.3597753>

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the Owner/Author.