Statistica Sinica **37** (2027), 1-22
doi:https://doi.org/10.5705/ss.202024.0153

TESTS ON DYNAMIC RANKING

Nan Lu[1,2], Jian Shi*[1,2], Xin-Yu Tian[3] and Kai Song[1]

[1]*Chinese Academy of Sciences,* [2]*University of Chinese Academy of Sciences and* [3]*University of Minnesota*

Abstract: In this paper, we investigate the dynamic Bradley–Terry model, a highly acclaimed statistical ranking model, and tackle several crucial inference problems related to score functions and rank properties. Specifically, we tackle the test problems of score function variation and pairwise similarity, providing valuable insights for model determination and simplification. We derive asymptotic null distributions for the proposed test statistics and prove the tests' consistency. Furthermore, we introduce a novel confidence band of dynamic rank and establish an innovative and generally applicable test framework for dynamic ranking properties. To overcome the conservativeness issue brought by the supreme form statistics, we introduce a novel approach based on signed score difference statistics for ranking inferences. We present theoretical guarantees for the proposed scheme. Numerical simulations validate the theories and demonstrate the satisfactory performance of our methods. The proposed methods are applied to a real dataset, yielding insightful results.

Key words and phrases: Bradley–Terry model, combinatorial inference, dynamic ranking property, time-varying test, uncertainty quantification.

1. Introduction

Ranking problems are widely studied in various disciplines, including information retrieval (Cambazoglu et al., 2010), computational biology (Duh and Kirchhoff, 2008), recommender systems (Lv et al., 2011), and sports games (Bong et al., 2020), among others. These problems have garnered considerable attention and have been extensively explored across different fields (Wang et al., 2013; Chen et al., 2019; Liu, Fang and Lu, 2023). The Bradley–Terry (BT) model (Bradley and Terry, 1952) is one of the most widely acclaimed statistical ranking models.

Though there have been tremendous studies on the BT model, such as Simons and Yao (1999), Hunter (2004), Huang, Weng and Lin (2006), Yan, Yang and Xu (2012), Han et al. (2020), and Gao, Shen and Zhang (2023), most of them focus on the static case. However, in reality, data is often collected with timestamps, and it is more reasonable to assume that audience preferences for movies change over time, as well as the potential ability of a basketball team is time-varying. Some studies on the dynamic BT model

*Corresponding author. E-mail: jshi@iss.ac.cn.