Ordinal Regression with Fenton-Wilkinson Order Statistics: A Case Study of an Orienteering Race

Authors : Joonas Pääkkönen

Abstract : In sports, individuals and teams are typically interested in final rankings. Final results, such as times or distances, dictate these rankings, also known as places. Places can be further associated with ordered random variables, commonly referred to as order statistics. In this work, we introduce a simple, yet accurate order statistical ordinal regression function that predicts relay race places with changeover-times. We call this function the Fenton-Wilkinson Order Statistics model. This model is built on the following educated assumption: individual leg-times follow log-normal distributions. Moreover, our key idea is to utilize Fenton-Wilkinson approximations of changeover-times alongside an estimator for the total number of teams as in the notorious German tank problem. This original place regression function is sigmoidal and thus correctly predicts the existence of a small number of elite teams that significantly outperform the rest of the teams. Our model also describes how place increases linearly with changeover-time at the inflection point of the log-normal distribution function. With real-world data from Jukola 2019, a massive orienteering relay race, the model is shown to be highly accurate even when the size of the training set is only 5% of the whole data set. Numerical results also show that our model exhibits smaller place prediction root-mean-square-errors than linear regression, mord regression and Gaussian process regression.

Keywords : Fenton-Wilkinson approximation, German tank problem, log-normal distribution, order statistics, ordinal regression, orienteering, sports analytics, sports modeling

Conference Title : ICSDMS 2020 : International Conference on Sports Data Mining and Statistics
Conference Location : Zurich, Switzerland
Conference Dates : July 27-28, 2020