

COMBINING P -VALUES USING HEAVY-TAILED DISTRIBUTIONS

Junsik Kim and Junyong Park*

Duksung Women's University and Seoul National University

Abstract: Combining individual p -values to handle large scale inferences or to aggregate results of different studies is one of major interest in meta-analysis which has been traditionally based on independent p -values. In contrast to combining methods that are constructed when p -values are independent, recently proposed combinations of p -values transformed into heavy-tailed distribution are known to be robust to the dependence structure of p -values. In this paper, we investigate theoretical properties of combining p -value methods for different heaviness of transformation under a wider class of correlation structures compared to existing studies from the viewpoint of controlling Type I error and obtaining powers. We also investigate relationships between harmonic mean type combination methods and combining methods that use transformation of p -values into stable distribution including Cauchy and Lévy combination methods. We provide extensive numerical studies supporting theoretical results. We also apply these p -value combining methods to real example of Crohn's disease data and present some idea on how to validate these methods.

Key words and phrases: Cauchy transformation, combining p -values, GWAS, Lévy transformation, meta-analysis.

1. Introduction

Meta-analysis has been used in various fields as a statistical technique to draw more reliable conclusions by combining the results of different studies or experiments. Most of research has focused on how to aggregate p -values generated from the assumption that all experiments or studies are independent. Among the methods for combining such independent p -values, well known methods including Fisher's method in Fisher (1934) and Stouffer's method in Stouffer et al. (1949) are based on aggregating some transformed p -values leading to chi-square distribution and normal distribution, respectively. In particular, the distribution of the sum of independent p -values or other forms of statistics needs to be derived relatively easily for practical use. For example, if the p -value is transformed into an infinitely divisible random variable, the distribution of the sum of those transformed p -values is the same type of distribution as the transformed distribution as in the case of Fisher's method and Stouffer's method.

*Corresponding author. E-mail: junyongpark@snu.ac.kr