

Boosting heritability: estimating the genetic component of phenotypic variation with multiple sample splitting

The Tien Mai⁽¹⁾_{*}, Paul Turner^{(3),(4)} and Jukka Corander^{(1),(2)}

⁽¹⁾ Oslo Centre for Biostatistics and Epidemiology, Department of Biostatistics, University of Oslo, Norway.

⁽²⁾ Department of Mathematics and Statistics, University of Helsinki, Finland.

⁽³⁾ Cambodia-Oxford Medical Research Unit, Angkor Hospital for Children, Siem Reap, Cambodia.

⁽⁴⁾ Centre for Tropical Medicine and Global Health, Nuffield Department of Medicine, University of Oxford, UK.

Abstract

Background: Heritability is a central measure in genetics quantifying how much of the variability observed in a trait is attributable to genetic differences. Existing methods for estimating heritability are most often based on random-effect models, typically for computational reasons. The alternative of using a fixed-effect model has received much more limited attention in the literature.

Results: In this paper, we propose a generic strategy for heritability inference, termed as “*boosting heritability*”, by combining the advantageous features of different recent methods to produce an estimate of the heritability with a high-dimensional linear model. Boosting heritability uses in particular a multiple sample splitting strategy which leads in general to a stable and accurate estimate. We use both simulated data and real antibiotic resistance data from a major human pathogen, *Sptreptococcus pneumoniae*, to demonstrate the attractive features of our inference strategy.

Conclusions: Boosting is shown to offer a reliable and practically useful tool for inference about heritability.

Keywords: *antimicrobial resistance, boosting, heritability, linear model.*

*Corresponding author: t.t.mai@medisin.uio.no