

Using IBS-based haseman-elston regression to estimate the heritability of complex characteristics from genome-wide association studies

Aldous Huxley

Huxley A . Using IBS-based haseman-elston regression to estimate the heritability of complex characteristics from genome-wide association studies. *J. Clin. Genet. Genom.* 2022;5(3):1-2.

ABSTRACT

Statistical genetics' main objective is to investigate the heritability of complicated traits. The variance component approaches are favourable when estimating heritability using markers, compared to other previously proposed methods. The best suitable heritability estimator model is frequently ambiguous since data from genome-wide association studies, in which genetic architecture is frequently unknown, have a high degree of dimension. Initially primarily suggested for linkage research, the Haseman-Elston regression is a variance component technique. This study, however, provides a theoretical foundation for a modified HE that models linkage disequilibrium for a quantitative trait and can, thus, be used for GWAS. We applied the IBS-based HE regression to single-marker

association studies (scenario I) and evaluated the variance component by employing multiple markers after substituting identical by descent scores with identity by state scores (scenario II). In scenario II, we go over the conditions under which the HE regression and the mixed linear model are similar; the difference between these two approaches is noticeable when an additive variance-covariance component is present. In a follow-up simulation analysis, we discovered that the IBS-based HE regression offered a more accurate heritability estimate than the mixed linear model when we used it to conduct case-control studies.

Key Words: Heritability; Linkage; Statistical genetics ; Single Marker; Primo-colonization

INTRODUCTION

Small sample sizes, an underrepresented variation spectrum, poor experimental design, and incorrect methodological assumptions are only a few of the causes of so-called "missing heritability." Estimating heritability is challenging due to the high dimensionality of genome-wide association study data, where N represents the number of people and M represents the number of markers. A strict p -value threshold of 10^{-8} , for example, may not be able to catch variations associated with a minor effect if the statistical power is insufficient. The mixed linear model, which relies on the genetic link between people inferred from single nucleotide polymorphism markers rather than fitting hundreds of thousands of markers together, can be used to partially get around this problem. However, a weighted genetic relationship matrix under various genetic architectures, which is frequently unknown, has been offered as a way to adapt an estimator under genetic architecture. Speed's ad-hoc weighing approach depends on the genetic architecture and does not frequently perform better than basic weight methods when compared,

as shown in extensive empirical data investigations. There should be criteria created to support the model used to estimate heritability because the genetic architecture, such as the link between variant frequency and variant effect, is frequently unclear. As more samples are gathered to examine diseases, many research projects eventually switch to a case-control methodology for GWAS. Scale transformation is required due to the case-control study ascertainment. Without scale translation, the observed scale's heritability can exceed 1, making the estimated heritability meaningless because it is not indicative of the liability scale's heredity, which is easier to interpret for illness data. According to Sang Hong Lee's paper, an equation that converts heritability from the observable scale to the liability scale has been proposed. This equation was examined using the infinitesimal model, in which there are an unlimited number of accidental loci. The question of whether Hong23 performs well for mixed linear model estimates if the infinitesimal model is rejected arises since, in reality, illness loci are typically very constrained for many diseases.

The heritability calculated by variance component approaches used

Editorial Office, *Journal of Clinical Genetics and Genomics*, UK

Correspondence: Aldous Huxley, *Journal of Clinical Genetics and Genomics*, UK, Email: aldoushuxley@gmail.com

Received:06-June-2022, Manuscript No. PULJCGG-22-5963; Editor assigned:08-June-2022, Pre-QC No. PULJCGG-22-5963 (PQ); Reviewed:- 12-June-2022, QC No. PULJCGG-22-5963 (Q); Revised:15-June-2022; Manuscript No. PULJCGG-22-5963 (R); Published:-20-June-2022, DOI: 10.37532/puljcg.22.5(3) 1-2



This open-access article is distributed under the terms of the Creative Commons Attribution Non-Commercial License (CC BY-NC) (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits reuse, distribution and reproduction of the article, provided that the original work is properly cited and the reuse is restricted to noncommercial purposes. For commercial reuse, contact reprints@pulsus.com

up to this point in mixed linear models is the subject of all of the aforementioned worries. A prominent method for estimating variance components is the Haseman-Elston regression. However, the HE regression, a well-known method for linkage studies that leverages identity by descent cores, appears to be a rusted weapon in the genomics analysis arsenal of the GWAS era. This is so because the HE regression uses IBD-measured relatedness but not state-measured identity. Although IBS has been used for linkage analysis, such as in the affected-pedigree-member design, its performance depends heavily on marker polymorphisms and might result in a significant number of false positives if ad-hoc weighting algorithms or the wrong frequencies are used. IBS, a notion that was underutilised throughout the linkage era, is neither used in the original HE regression framework nor is it well suited to linkage investigations. Recently, a novel idea that uses like-standardized IBS has opened up a new method for determining the genetic relatedness of unrelated people. It is unclear if the IBS score can be applied to the HE regression for unrelated individuals because it is similar to the traditional IBD score. We used the HE regression to conduct association studies for GWAS data in this study, replacing IBD scores with standardised IBS ratings. This article lays the theoretical groundwork for applying the HE regression for GWAS, presuming random mating, biallelic loci, and additive genetic effects only on the genetic architecture of quantitative trait loci underlying a complex variable. Regression coefficients for two general scenarios were calculated, and their interpretations are relevant to genetics. In case I, a marker in linkage disequilibrium with a QTL was used to determine the IBS score. Due to this, HE regression could be used as a technique for single-marker GWAS. In case II, the IBS score was calculated based on a number of markers, each of which might be in LD with a number of QTLs. This made it possible to estimate the variance component marked by markers using the HE regression.