広島統計談話会

Hiroshima Statistics Study Group

第 312 回談話会を下記のように開催致しますので

御参集下さいますようご案内申し上げます。

You are cordially invited to the 312th meeting as scheduled below.

日　時：　2018 年 9 月 7 日（金）15:00 –

Date　：　September 7th, 2018 (Fri) 15:00 –

場　所：　放射線影響研究所　講堂

Place　：　RERF Auditorium

演　者：　植木 優夫 博士（理化学研究所 革新知能統合研究センター 研究員）

Speaker　：　Masao Ueki, Ph.D.
Researcher
RIKEN Center for Advanced Intelligence Project

演　題：　「双方向グラフ上の最短経路を用いた遺伝的関連性の検出」

Title　：　“Detecting genetic association through shortest paths in a bidirected graph”

要　約：

Summary:

Genome-wide association studies (GWASs) commonly use marginal association tests for each single-nucleotide polymorphism (SNP). Because these tests treat SNPs as independent, their power will be suboptimal for detecting SNPs hidden by linkage disequilibrium (LD). One way to improve power is to use a multiple regression model.

However, the large number of SNPs preclude simultaneous fitting with multiple regression, and subset regression is infeasible because of an exorbitant number of candidate subsets. We therefore propose a new method for detecting hidden SNPs having significant yet weak marginal association in a multiple regression model. Our method begins by constructing a bidirected graph locally around each SNP that demonstrates a moderately sized marginal association signal, the focal SNPs. Vertexes correspond to SNPs, and adjacency between vertexes is defined by an LD measure. Subsequently, the method collects from each graph all shortest paths to the focal SNP. Finally, for each shortest path the method fits a multiple regression model to all the SNPs lying in the path and tests the significance of the regression coefficient corresponding to the terminal SNP in the path. Simulation studies show that the proposed method can detect susceptibility SNPs hidden by LD that go undetected with marginal association testing or with existing multivariate methods. When applied to real GWAS data from the Alzheimer's Disease Neuroimaging Initiative (ADNI), our method detected two groups of SNPs: one in a region containing the apolipoprotein E (APOE) gene, and another in a region close to the semaphorin 5A (SEMA5A) gene.