S592 Literature Report 2

Optimal tests for rare variant effcts in sequencing association studies

Nick Sun

March 3, 2019

Main Contributions

Rare variant association tests are statistical procedures whose goal is to detect the effects of rare alleles variants on phenotypes of interest such as serum triglyceride levels. Since these genetic variants are very rare, it is difficult to test these effects using traditional statistical methods. Researchers often employ specialized techniques to test these effects. There are currently two predominantly used tests in these scenarios: weighted burden tests and sequence kernel association tests.

Weighted burden tests are designed around treating each rare variant as having an effect in the same direction (no variants counteract each other). In addition, it is also assumed that each rare variant has the same magnitude on the phenotype. These tests have good power under these assumptions, but in practice with many rare variants, it is likely that these assumptions are not met. As a result, the power and false discovery rate of WBT tests suffer.

Sequence kernel association tests, or SKAT tests, by converse allow rare variants to have different effect directions but less power than weighted burden tests when most of the effects are in the same direction.

The authors propose a new novel technique called SKAT-O for finding the optimal test among a family of proposed tests. This family of tests is constructed as a convex combination of WBT and SKAT tests. Additionally, the authors provide analytic formulas for appropriate sample size and power calculations for SKAT tests.

SKAT-O overview

Both the WBT and SKAT tests have the following set up which SKAT-O extends. The response y_i follows an distribution from an exponential family with a link function $g(\mu_i) = X_i\alpha + G_i\beta$ where α and β are vectors of regression coefficients for covariates and the rare variants.

WBT proposes using the minor allele frequency (MAF) as a weight of the j rare variants and restructuring the model so that we only have to test one parameter β_0 for significance. This reduces the degrees of freedom which is necessary when we are testing a large number of variants. Conversely, SKAT reduces the degrees of freedom by assuming that each β_j is a random effect with variance ψ. The test for significant β then becomes equivalent to testing whether $\psi = 0$.

The new family of tests for SKAT-O corrects for the shortcomings of SKAT by allowing a correlation factor between the β parameters. When this correlation factor is included in the kernel function, this helps SKAT perform comparably to the WBT test when a large percentage of the variants in the targeted region are associated with the phenotype of interest in the same direction. Therefore, the construction of the proposed family of tests relies on heavily on a parameter ρ. This ρ parameter represents the correlation between the individual β_i in the model.

The procedure for finding ρ is a simple grid search that tests a lot of different values of ρ. The resulting test statistic of SKAT-O is the minimum p-value of all these tested ρ. The p-values are individually calculated from Q_ρ, a test statistic that approximately follows a mixture of χ^2 random variables with an additional variance component ζ which incorporates the additional uncertainty of having a dynamic ρ. This Q_ρ was demonstrated to be a convex combination of the test statistics of the SKAT test and the WBT test: $Q_\rho = (1-\rho)Q_{\text{SKAT}} + \rho Q_{\text{burden}}$

Additionally to the new SKAT-O tests, the authors also provide novel formulas for calculating theoretical power of the SKAT test using an "optimal" value of ρ found using a simple formula $\rho = p_1^2(2p_2 = 1)^2$ where p_1 is the proportion of nonzero βs and p_2 is the proportion of the nonzero βs that are positive.

Examples

The authors have several examples: a simulation study and an example using real data. From the simulation study, they were able to demonstrate that all tests (WBT, SKAT, and SKAT-O) were relatively the same in terms of exactness and SKAT-O's power was close to WBT when the rare variants were mostly causal and in the same direction as well as close to SKAT when the proportion of causal rare variants was low. The real data example consisted of examining known causal rare variants which affect serum triglyceride levels in the blood. Again, SKAT-O was shown to perform relatively well compared to the other tests.

The final argument the authors made is that SKAT-O is not as good as either test when the specific circumstances for which WBT and SKAT were designed are met. However, SKAT-O's power is still fairly close to these tests and is more robust than either test by itself.

Caveats and critical remarks

The main caveat is that SKAT-O falls in between the SKAT and WBT tests. It is akin to having a person who is somewhat skilled at two separate situations as opposed to having two people who are very specialized and excel at individual situations. The real data example demonstrated a clear power tradeoff that was made for additional robustness in SKAT-O. Even when the optimal ρ was found to be 0 or 1, the power of the SKAT-O test was not as high as the SKAT or WBT test respectively even though theoretically the tests should have been equivalently. This power sacrifice is a result of having to empirically search for an optimal ρ and is an intrinsic weakness in the test. Therefore, a natural question is if it is worth using SKAT-O when researchers have a good informed hypothesis about the causal variants in question and which direction the effects should be in.

Another caveat was that the ρ calculated for the theoretical power calculations is likely to be different than the ρ found using a grid search for the SKAT-O test. Therefore, the formulas cannot be used exactly for the SKAT-O test. The authors assuaged this by saying that in practice, the difference in empirical power and theoretical power is small and demonstrated this using the simulated data.