In a variance component model, the genetic effects for rare variants from each nucleotide site are modeled as random effects with mean 0 and variance $\eta^2$, i.e. $dP(\tilde{\beta}_2;0,\eta^2)$. According to the work by Lin [1], the variance component score statistics can be calculated based upon the marginal likelihood,

$$p\left(\bar{Y}_2|A_i,\bar{X}_i;\eta,\tau_1,\tau_2,\mu(Y_{i1}),\mu(Y_{i2})\right) = \prod_i p\left(\bar{Y}_i|A_i,\bar{X}_i;\tilde{\beta}_2,\tau_1,\tau_2,\mu(Y_{i1}),\mu(Y_{i2})\right) dP\left(\tilde{\beta}_2;0,\eta^2\right) \quad (S5.1)$$

However, the calculation using (S5.1) is not straightforward. Instead, the following theorem can be applied:

**Theorem:** The variance component score statistics calculated using (S5.1) can be equivalently constructed based upon the conditional likelihood, i.e.

$$p\left(\bar{Y}_2|\bar{Y}_i,A_i,\bar{X}_i;\eta,\tau_1,\tau_2,\mu(Y_{i1}),\mu(Y_{i2})\right) = \int \prod_i p\left(Y_{i2}|Y_{i1},A_i,\bar{X}_i;\tilde{\beta}_2,\tau_1,\tau_2,\mu(Y_{i1}),\mu(Y_{i2})\right) dP\left(\tilde{\beta}_2;0,\eta^2\right) \quad (S5.2)$$

Therefore the variance component score statistic has the form

$$S = \left(\bar{Y}_2 - \hat{\mu}(\bar{Y}_2)\right)^T Ker\left(\bar{X},\bar{X}\right) \left(\bar{Y}_2 - \hat{\mu}(\bar{Y}_2)\right)$$

where $Ker\left(\bar{X}_i,\bar{X}_j\right)$ is the kernel function for comparing two multi-site genotypes, and $\hat{\mu}(\bar{Y}_i)$, $\hat{\mu}(\bar{Y}_2)$ are the estimated mean primary and secondary trait values under the null model.

**Proof:** In order to prove the theorem, we first factorize the probability

$$p\left(Y_{i1},Y_{i2}|A_i,\bar{X}_i;\tilde{\beta}_2,\tau_1,\tau_2,\mu(Y_{i1}),\mu(Y_{i2})\right), \text{ i.e.}$$

$$p\left(Y_{i2}|Y_{i1},A_i,\bar{X}_i;\tilde{\beta}_2,\mu(Y_{i1}),\mu(Y_{i2})\right) = p\left(Y_{i2}|A_i,\bar{X}_i;\tilde{\beta}_2,\tau_1,\tau_2,\mu(Y_{i1}),\mu(Y_{i2})\right)$$

$$\times p\left(Y_{i1}|A_i,\bar{X}_i;\tau_1,\tau_2,\mu(Y_{i1}),\mu(Y_{i2})\right).$$

The term $p\left(Y_{i1}|A_i,\bar{X}_i;\tau_1,\tau_2,\mu(Y_{i1}),\mu(Y_{i2})\right)$ does not contain the parameters $\tilde{\beta}_2$, and therefore can be taken out of the integral in (S5.1), i.e.
If the sample ascertainment is only based upon the primary trait, using the same argument as in equations (8) and (9), we can obtain

\[
p(Y_{i2}|Y_{i1}, A, \tilde{X}; \tau_2, \mu(Y_{i1}), \mu(Y_{i2})) = p(Y_{i2}|Y_{i1}, \tilde{X}; \tau_3, \mu(Y_{i1}), \mu(Y_{i2}))
\]

and

\[
p(Y_{i2}|Y_{i1}, A, \tilde{X}; \tau_1, \tau_2, \mu(Y_{i1}), \mu(Y_{i2})) = p(Y_{i2}|Y_{i1}, \tilde{X}; \tau_1, \tau_2, \mu(Y_{i1}), \mu(Y_{i2}))
\]

Therefore, the variance component score statistics can be equivalently calculated from

\[
p(\tilde{Y}_2|Y_{i1}, \tilde{X}; \tau_1, \tau_2, \tau_3, \mu(Y_{i1}), \mu(Y_{i2})) = \prod_{i} p(Y_{i2}|Y_{i1}, \tilde{X}; \tau_1, \tau_2, \tau_3, \mu(Y_{i1}), \mu(Y_{i2}))dP(\beta; 0, \eta^2).
\]

Given that \(\hat{\mu}(Y_{i1}), \hat{\mu}(Y_{i2})\) are consistent estimators under the null hypothesis, by Slutsky’s theorem [2], \(Y_{i2} - \hat{\mu}(Y_{i2})\) is approximately normally distributed conditional on \(Y_{i1}, \tilde{X}_i\), i.e.

\[
Y_{i2} - \hat{\mu}(Y_{i2}) \sim \mathcal{N}(0, \tau^2_2)
\]

Therefore, according to Wu et al [3], the SKAT statistic is given by

\[
S = (\tilde{Y}_2 - \hat{\mu}(\tilde{Y}_2))^\top Ker(\tilde{X}, \tilde{X}) \left(\tilde{Y}_2 - \hat{\mu}(\tilde{Y}_2)\right).
\]

The SKAT statistic follows a mixture chi-square distribution with the mixture proportions being the eigen values for \(Ker(\tilde{X}, \tilde{X})\var(\tilde{Y}_2 - \hat{\mu}(\tilde{Y}_2))\).

Reference: