Untitled

PHS 7095, Spring 2024 - Final Report

Kline Dubose, Haojia Li

05/05/2024

1 Introduction

- Non-standard because the parameter under the null hypothesis is on the boundary of the parameter space and the response variables are not independent under the alternative.
- Special case of variance component of familial relatedness in linear mixed models.

2 Likelihood Ratio Tests in Linear Mixed Models with One Variance Component

Crainiceanu and Ruppert (2004) derived both the finite sample distribution and the asymptotic distribution of the LRT and RLRT statistics for testing the null hypothesis that the variance component is 0 in a linear mixed model (LMM) with one variance component. They used weak assumptions on eigenvalues of certain design matrices, and released the hypothesis of i.i.d. data in the restrictive assumptions by Self and Liang (1987), that the response variable vector can be partitioned into independent and identically distributed subvectors and the number of independent subvectors tends to ∞.

2.1 Model and hypotheses setup

Consider an LMM with one variance component

$$\mathbf{Y} = X\beta + Z\mathbf{b} + \epsilon, \ E \begin{pmatrix} \mathbf{b} \\ \epsilon \end{pmatrix} = \begin{pmatrix} \mathbf{0}_K \\ \mathbf{0}_n \end{pmatrix}, \ \text{cov} \begin{pmatrix} \mathbf{b} \\ \epsilon \end{pmatrix} = \begin{pmatrix} \sigma_b^2 \Sigma & \mathbf{0} \\ \mathbf{0} & \sigma_\epsilon^2 I_n \end{pmatrix}, \tag{1}$$

where,

- $\mathbf{Y}$ is the $n \times 1$ vector of observations,
- X is the $n \times p$ design matrix for the fixed effects,
- Z is the $n \times K$ design matrix for the random effects,
- β is a p-dimensional vector of fixed effects parameters,
- $\mathbf{b}$ is a K-dimensional vector of random effects,
- $(\mathbf{b}, \epsilon)$ has a normal distribution.

Under these conditions it follows that

$$E(\mathbf{Y}) = X\beta,$$
$$\text{var}(\mathbf{Y}) = \sigma_\epsilon^2 V_\lambda$$

where

- $\lambda = \sigma_b^2/\sigma_\epsilon^2$, which can be considered a signal-to-noise ratio,
- $V_\lambda = I_n + \lambda Z\Sigma Z'$.

In this context, $\sigma_b^2 = 0$ if and only if $\lambda = 0$, and the parameter space for λ is $[0, \infty)$.

The authors proposed a general form of the null hypothesis (H_0) and alternative hypothesis (H_A):

$$H_0 : \beta_{p+1-q} = \beta_{p+1-q}^0, ..., \beta_p = \beta_p^0, \qquad \sigma_b^2 = 0 \text{ (or equivalently } \lambda = 0)$$
$$H_A : \beta_{p+1-q} \neq \beta_{p+1-q}^0, ..., \beta_p \neq \beta_p^0, \qquad \sigma_b^2 > 0 \text{ (or equivalently } \lambda > 0)$$

where the q denotes the number of fixed effects parameters constrained under H_0. An application of this method with $q > 0$ is using LRTs and RLRTs for testing a polynomial fit against a general alternative described by penalized splines (P-splines), in Section 5 of the paper.

However, we will only focus on the particular case of $q = 0$:

$$H_0 : \sigma_b^2 = 0 \ (\lambda = 0) \quad \text{vs.} \quad H_A : \sigma_b^2 > 0 \ (\lambda > 0)$$

2.2 Finite sample distribution of LRT and RLRT

2.2.1 LRT (Theorem 1)

Twice the log-likelihood ratio statistic for model (1) is

$$2\log L(\beta, \lambda, \sigma_e^2) = -\log(\sigma_e^2) - \log|V_\lambda| - \frac{(Y - X\beta)'V_\lambda^{-1}(Y - X\beta)}{\sigma_e^2} \tag{2}$$

and the LRT is defined as

$$LRT_n = 2\sup_{H_A}\{L(\beta, \lambda, \sigma_\epsilon^2)\} - 2\sup_{H_0}\{L(\beta, \lambda, \sigma_\epsilon^2)\}$$

Under the alternative hypothesis, by fixing λ and solving the first-order maximum conditions for β and σ_ϵ^2, we obtain the profile likelihood estimates

$$\hat{\beta}(\lambda) = (X'V_\lambda^{-1}X)^{-1}X'V_\lambda^{-1}\mathbf{Y}, \quad \hat{\sigma}_\epsilon^2(\lambda) = \frac{(\mathbf{Y} - X\hat{\beta}(\lambda))'V_\lambda^{-1}(\mathbf{Y} - X\hat{\beta}(\lambda))}{n}.$$

Plugging these estimates back into the log-likelihood function, we obtain the profile log-likelihood function

$$L^{K,n}(\lambda) = -\log|V_\lambda| - n\log(\mathbf{Y}'P_\lambda'V_\lambda^{-1}P_\lambda\mathbf{Y})$$

where $P_\lambda = I_n - X(X'V_\lambda^{-1}X)^{-1}X'V_\lambda^{-1}$ and $P_0 = I_n - X(X'X)^{-1}X'$ under the null hypothesis when the model becomes a standard linear regression.

The LRT statistic can be written as

$$LRT_n = \sup_{\lambda \geq 0}\{n\log(\mathbf{Y}'P_0\mathbf{Y}) - n\log(\mathbf{Y}'P_\lambda'V_\lambda^{-1}P_\lambda\mathbf{Y}) - \log|V_\lambda|\}$$

The following theorem gives the spectral decomposition of the LRT_n statistic for testing the null hypothesis (2) against the alternative hypothesis (3). Define

$$f_n(\lambda) = n\log\left\{1 + \frac{N_n(\lambda)}{D_n(\lambda)}\right\} - \sum_{s=1}^{K}\log(1 + \lambda\xi_{s,n}),$$

where $N_n(\lambda)$ and $D_n(\lambda)$ are defined in theorem 1.

Theorem 1. If $\mu_{s,n}$ and $\xi_{s,n}$ are the K eigenvalues of the $K \times K$ matrices $\Sigma^{1/2}Z^TP_0Z\Sigma^{1/2}$ and $\Sigma^{1/2}Z^TZ\Sigma^{1/2}$ respectively, where $P_0 = I_n - X(X^TX)^{-1}X^T$, then

$$LRT_n = n\left(1 + \sum_{s=1}^{q}\frac{u_s^2}{w_s^2} - \frac{n-p}{\sum_{s=1}^{n-p}w_s^2} + \sup_{\lambda \geq 0}f_n(\lambda)\right),$$

where u_s for $s = 1,\dots,K$ and w_s for $s = 1,\dots,n-p$ are independent $N(0,1)$, the notation $\Rightarrow$ denotes equality in distribution and

$$N_n(\lambda) = \sum_{s=1}^{K}\frac{\lambda\mu_{s,n}}{1 + \lambda\mu_{s,n}},$$

$$D_n(\lambda) = \sum_{s=1}^{K}\frac{w_s^2}{1 + \lambda\mu_{s,n}} + \sum_{s=K+1}^{n-p}w_s^2.$$

The spectral representations (7) and (9) below can be used to compute the probability mass at 0 of the LRT and RLRT statistic for testing that the variance component is 0 with no constraints on the fixed effects ($q = 0$ and $\sigma_b^2 = 0$). The first-order condition for $f_n(\lambda)$ to have a local maximum at $\lambda = 0$ is

$$\left.\frac{\partial f_n(\lambda)}{\partial\lambda}\right|_{\lambda=0} \leq 0.$$

It follows that the probability of having a local maximum of the profile likelihood at $\lambda = 0$ is

$$\mathrm{pr}\left(\frac{\sum_{s=1}^{K}\mu_{s,n}w_s^2}{\sum_{s=1}^{n-p}w_s^2} \leq \frac{1}{n}\sum_{s=1}^{K}\xi_{s,n}\right). \tag{8}$$

3

2.2.2 Proof of Theorem 1

We partition first the vector of fixed effect parameters $\beta = (\beta_1^T \beta_2^T)^T$, where $\beta_2 = \beta_{20}$ are the known fixed effect parameters under the null hypothesis. We can also partition the matrix of fixed effects $X = (X_1 X_2)$ corresponding to the partition of β. Using the notations in the paper, the LRT statistic can be written as:

$$LRT_n = \sup_{\lambda \geq 0} \left\{ n \log \left(\frac{y^T P_0 y}{y^T P_\lambda^{-1} y} \right) - n \log(|V_\lambda|) + n \log(|S_1|) \right\}$$

The first part of the right hand side of equation (1) corresponds to testing for the zero variance of random effects while the second part corresponds to testing for the fixed effects. One can easily show that $\log|V_\lambda| = -\sum_{s=1}^{K} \log(1 + \lambda \xi_{s,n})$. Also, from Kuo, 1999 and Patterson and Thompson, 1971 there exists an $n \times (n-p)$ matrix W such that $W^T W = I_{n-p}, W^T P_0 = P_0, W^T V_\lambda = \text{diag}(1 + \lambda \mu_{s,n})$, and

$$y^T P_\lambda^{-1} y = y^T W \text{diag}(1 + \lambda \mu_{s,n})^{-1} W^T y.$$

Denote by $w = W^T y / \sigma_c$ and note that under the null hypothesis

$$E[w] = (W^T X_1 \beta_1 + W^T X_2 \beta_2)/\sigma_c, \quad \text{Cov}[w] = I_{n-p}.$$

We now show that $E[w] = 0$. Denote by $A = W^T X$ and observe that $WA = P_0 X = 0$, and that $W^T W A = 0$. This shows that $A = 0$, that $W^T X_1 = 0$, and that $W^T X_2 = 0$. It now follows that $w = (w_1, ..., w_{n-p})$ is an $n - p$ dimensional random vector with i.i.d. $N(0,1)$ components. Putting all these together it follows that:

$$L_n(\lambda) = -n \log \left(\sigma_c^2 \left(\sum_{s=1}^{K} \frac{w_s^2}{1 + \lambda \mu_{s,n}} + \sum_{s=K+1}^{n-p} w_s^2 \right) \right) - \sum_{s=1}^{K} \log(1 + \lambda \xi_{s,n}),$$

where we used the fact that at most K eigenvalues $\mu_{s,n}$ and $\xi_{s,n}$ are not zero. In particular

$$L_n(0) = -n \log \left(\sigma_c^2 \left(\sum_{s=1}^{n-p} w_s^2 \right) \right),$$

which is a standard result in regression analysis. Therefore we can write

$$L_n(\lambda) - L_n(0) = n \log \left(1 + U_n(\lambda) \right),$$

where $U_n(\lambda) = N_n(\lambda)/D_n(\lambda)$ and

$$N_n(\lambda) = \sum_{s=1}^{K} \frac{\lambda \mu_{s,n} w_s^2}{1 + \lambda \mu_{s,n}}, \quad D_n(\lambda) = \sum_{s=1}^{K} \frac{w_s^2}{1 + \lambda \mu_{s,n}} + \sum_{s=K+1}^{n-p} w_s^2.$$

2.2.3 RLRT

Residual, or restricted, maximum likelihood (REML) was introduced by Patterson and Thompson (1971) to take into account the loss in degrees of freedom due to estimation of β-parameters. REML consists of maximizing the likelihood function that is associated with $n - p$ linearly independent contrasts, and the log-likelihood function for any such set of contrasts differs by no more than an additive constant (Harville, 1977). Twice the restricted profile log-likelihood function is

$$2l_n^*(\lambda) = -\log|\mathbf{V}_\lambda| - \log|\mathbf{X}^T\mathbf{V}_\lambda^{-1}\mathbf{X}| - (n-p)\log(\mathbf{Y}^T\mathbf{P}_\lambda\mathbf{V}_\lambda^{-1}\mathbf{P}_\lambda\mathbf{Y}).$$

The RLRT_n is defined like the LRT_n by using the restricted likelihood instead of the likelihood function. Because the RLRT_n uses the likelihood of residuals after fitting the fixed effects, the RLRT is appropriate for testing only when the fixed effects are the same under the null and alternative hypotheses. Therefore the RLRT_n will be used only when the number of fixed effects constrained under H_0 is $q = 0$ and we test for $\sigma_b^2 = 0$ only. Then, under the null hypothesis described in equation (4),

$$RLRT_n = \sup_{\lambda \geq 0} \left[(n-p)\log\left(1 + \frac{N_n(\lambda)}{D_n(\lambda)}\right) - \sum_{s=1}^{K} \log(1 + \lambda\mu_{s,n}) \right],$$

and the probability of having a local maximum at $\lambda = 0$ is

$$\text{pr}\left(\frac{\sum_{s=1}^{K} \mu_{s,n} w_s^2}{\sum_{s=1}^{n-p} w_s^2} \leq \frac{1}{n-p}\sum_{s=1}^{K} \mu_{s,n} \right),$$

where $w_1, \ldots, w_k$ are independent $N(0,1)$ random variables. The notation here is the same as in theorem 1. An efficient simulation algorithm for the finite sample distribution of RLRT_n can be easily obtained by direct analogy with the LRT_n case.

2.3 Asymptotic distribution of LRT and RLRT

This section presents the asymptotic distributions of the LRT_n and $RLRT_n$ statistics for testing the null hypotheses (2) and (4) respectively. Because the finite sample results in Section 2 depend essentially on the eigenvalues $\mu_{s,n}$ and $\xi_{s,n}$, we may expect to have a relationship between the asymptotic distributions of test statistics and the asymptotic behavior of these eigenvalues. The following theorem provides the formal description of this relationship.

Theorem 2. Assume that hypothesis H_0 in expression (2) is true. Suppose that there is an $\alpha \geq 0$ so that for every s the K eigenvalues $\mu_{s,n}$ and $\xi_{s,n}$ of matrices $\Sigma^{1/2}Z^TP_0Z\Sigma^{1/2}$ and $\Sigma^{1/2}Z^TZ\Sigma^{1/2}$ respectively satisfy $\lim_{n\to\infty} n^{-\alpha}\mu_{s,n} = \mu_s$ and $\lim_{n\to\infty} n^{-\alpha}\xi_{s,n} = \xi_s$, where not all ξ_s are 0. Then

$$LRT_n \Rightarrow \sum_{s=1}^{q} u_s^2 + \sup_{d \geq 0}\{LRT_\infty(d)\},$$

where $\Rightarrow$ denotes convergence in distribution.

where the two terms in the asymptotic distribution are independent, u_s, s are IID $N(0,1)$, and

$$LRT_\infty(d) = \sum_{s=1}^{K} \frac{d\mu_s}{1 + d\mu_s} w_s^2 - \sum_{s=1}^{K} \log(1 + d\xi_s).$$

Here $\Rightarrow$ denotes weak convergence. The proof of this result is based on the weak convergence of the profile LRT in the space $C[0, \infty)$ of continuous functions on $[0, \infty)$ and a continuous mapping theorem. The first part of the asymptotic distribution is a χ^2-distribution corresponding to testing for q fixed effects and the second part corresponds to testing for $\sigma_b^2 = 0$.

Asymptotic theory for a null hypothesis on the boundary of the parameter space developed for IID data suggests the asymptotic result

$$LRT_n \Rightarrow \sum_{s=1}^{q} u_s^2 + U^2,$$

where U^2 represents some asymptotically chi-squared distributed random variable.

Under the same assumptions as in theorem 2, if the null hypothesis H_0 in equation (4) is true then

$$RLRT_n \Rightarrow \sup_{d \geq 0}\{RLRT_\infty(d)\},$$

where

$$RLRT_\infty(d) = \sum_{s=1}^{K} \frac{d\mu_s}{1 + d\mu_s} w_s^2 - \sum_{s=1}^{K} \log(1 + d\mu_s).$$

If $q = 0$, then the LRT and RLRT statistics have probability mass at 0. Crainicceanu, Ruppert and Vogelsang (2003) showed that this mass can be very large for LRT_∞ and $RLRT_\infty$ and is equal to the null probability that $LRT_\infty(\cdot)$ and $RLRT_\infty(\cdot)$ have a global maximum at 0. The latter is well approximated by the null probability of having a local maximum at $\lambda = 0$. The first-order conditions for a local maximum at $d = 0$ are

$$\left.\frac{\partial LRT_\infty(d)}{\partial d}\right|_{d=0} \leq 0$$

or

$$\left.\frac{\partial RLRT_\infty(d)}{\partial d}\right|_{d=0} \leq 0.$$

Therefore, the probability of having a local maximum at $d = 0$ for $LRT_\infty(\cdot)$ or $RLRT_\infty(\cdot)$ is

$$\mathrm{pr}\left(\sum_{s=1}^{K} \mu_s w_s^2 \leq \sum_{s=1}^{K} \xi_s\right).$$

2.3.1   Proof of Theorem 2

We continue to use notations from the proof of theorem 1. For $R(x) = \log(1+x) - x$, $\lim_{x \to 0} R(x)/x = 0$ and $\lim_{x \to 0} R(x)/x^2 = -1/2$. Using the Taylor expansion around 0, $\log(1 + x) = x + R(x)$ and taking into account that $\sum_{s=K+1}^{n} w_s^2/n$ converges almost surely to 1 one obtains

$$n \log \left(1 + \frac{\sum_{s=1}^{q} u_s^2}{\sum_{s=1}^{n-p} w_s^2}\right) = \sum_{s=1}^{q} u_s^2 + V_n,$$

where V_n converges almost surely to 0. Denoting by $W_q = \sum_{s=1}^{q} u_s^2$ one obtains $LRT_n = \sup_{\lambda \geq 0} LRT_n(\lambda)$ where

$$LRT_n(\lambda) = n \log \left(1 + U_n(\lambda)\right) - \sum_{s=1}^{K} \log(1 + \lambda \xi_{s,n}) + W_q + V_n,$$

where $U_n(\lambda)$ is independent of W_q, and V_n converges almost surely to 0. Denote now by $f_n(d) = n \log \left(1 + U_n(n^{-\alpha} d)\right) - \sum_{s=1}^{K} \log(1 + dn^{-\alpha} \xi_{s,n}) + W_q$ and we will show that

$$\sup_{d \geq 0} f_n(d) \geq \sup_{d \geq 0} LRT_\infty(d) + W_q.$$

This proof consists of two steps:

1. Prove that $f_n(d)$ converges weakly to $LRT_\infty(\cdot) + W_q$ on the space $C[0, \infty)$ of continuous functions with support $[0, \infty)$.
2. Prove that a Continuous Mapping Theorem type result holds for the $\sup_{d \geq 0} f_n(d)$.

We show the weak convergence for any $C[0, M]$. Denote $f(d) = LRT_\infty(d) + W_q$, $n^{-\alpha} u_n = d$, $n^{-\alpha} \xi_{s,n} = \xi_{s,n}^\alpha$, etc. We establish the finite dimensional convergence of $f_n(d) \to f(d)$ and then we prove that $f_n(d)$ is a tight sequence in $C[0, M]$. Note that finite-dimensional convergence and tightness are sufficient to establish the convergence in $C[0, M]$.

To show finite dimensional convergence it is sufficient to show that for a fixed d the convergence is almost sure. Note that

$$n \log \left(1 + U_n(n^{-\alpha} d)\right) = n U_n(d) + n R(n^{-\alpha} d),$$

where $R(\cdot)$, $U_n(\cdot)$, $N_n(\cdot)$ and $D_n(\cdot)$ were defined earlier. It follows immediately that almost surely

$$\lim_{n \to \infty} n U_n(n^{-\alpha} d) = K \sum_{s=1}^{K} \frac{d \mu_s}{1 + d \mu_s} w_s^2.$$

Because $nR(U_n(n^{-\alpha} d)) = (n U_n(n^{-\alpha} d)) \times R(U_n(n^{-\alpha} d)/U_n(n^{-\alpha} d))$, it follows that $nR(U_n(n^{-\alpha} d))$ converges to zero almost surely ($\lim_{x \to 0} R(x)/x = 0$). Note that $\lim_{x \to 0} \log(1 + x) = x$ ensures that $\log(1 + d \xi_{s,n}^\alpha)$ also converges to zero for every fixed d. We proved that, for every fixed d, $f_n(d)$ converges almost surely to $LRT_\infty(d) + W_q$.

To show that $f_n(d)$ form a tight sequence it is sufficient to show that for every $\epsilon > 0$ and strictly positive, there exist $\delta = \delta(\epsilon), 0 < \delta < 1$ and $n_0 = n_0(\epsilon, \delta)$ such that for $n \geq n_0$

$$\frac{1}{\delta} P \left(\sup_{|u-s| < \delta} |f_n(u) - f_n(t)| \geq \epsilon \right) < \eta.$$

Observe first that

$$|f_n(u) - f_n(t)| \leq n \log \left(\frac{D_n(n^{-\alpha}t)}{D_n(n^{-\alpha}u)} \right) + \sum_{s=1}^{K} \left| \log \left(1 + u\xi_{s,n} \right) - \log \left(1 + t\xi_{s,n} \right) \right|.$$

And because $\log(1 + x) < x$ for every $x > 0$ we obtain the following inequalities

$$\log \left(\frac{D_n(n^{-\alpha}t)}{D_n(n^{-\alpha}u)} \right) \leq (s - t)C \sum_{s=K+1}^{n-p} w_s^2,$$

where $C > 0$ is a constant so that $n\xi_{s,n}(n - p - K) \leq C$ for every s and n. It follows that the following inequality holds

$$n \log \left(\frac{D_n(n^{-\alpha}t)}{D_n(n^{-\alpha}u)} \right) \leq (u - t)CKF_{K,n},$$

where $F_{K,n}$ is a random variable with an F distribution with $(K, n - p - K)$ degrees of freedom. Similarly,

$$\sum_{s=1}^{K} \left| \log \left(1 + u\xi_{s,n} \right) - \log \left(1 + t\xi_{s,n} \right) \right| \leq (u - t)CK.$$

We conclude that

$$P \left(\sup_{|u-s|<\delta} |f_n(u) - f_n(t)| \geq \epsilon \right) \leq P \left(F_{K,n} > C\delta^{-1} \right),$$

which bounds the probability as required.

It is sufficient to find δ, n_0 such that for every $n \geq n_0$ the c.d.f. $H_{K,n}$ of $F_{K,n}$ satisfies

$$1 - H_{K,n} \left(\frac{\epsilon}{C\delta^{-1}} - 1 \right) \leq \eta\delta.$$

If H_K is the c.d.f. of a χ^2 distribution with K degrees of freedom, then for every n, $\lim_{n \to \infty} H_{K,n}(x) = H_K(x)$. Because (using for example l'Hospital rule and the pdf of a χ^2 distribution with K degrees of freedom)

$$\lim_{x \to \infty} \left[1 - H_K \left(\frac{\epsilon}{C} - K \right) \right] \left(\frac{n}{2} \right)^{\frac{1}{2}} = 0,$$

one can find $\delta = \delta(\eta), \delta < 1$ with $\epsilon/C\delta^{-1} - K > 0$ such that $1 - H_K(\epsilon/C\delta^{-1} - K) \leq \sqrt{\eta}\delta/2$. Because of the convergence of $H_{K,n}$ to H_K, one can find $n_0 = n_0(\eta, \delta)$ so that for $n \geq n_0$ the following inequality holds

$$\left| H_{K,n} \left(\frac{\epsilon}{C\delta^{-1}} - K \right) - H_K \left(\frac{\epsilon}{C\delta^{-1}} - K \right) \right| \leq \frac{\sqrt{\eta}\delta}{2},$$

which finishes the proof of the inequality in equation (2). We conclude that $f_n(d)$ converges weakly to $f(d)$ on $C[0, M]$ for each M, and therefore on $C[0, \infty)$.

We want to show now that $\sup_{d \geq 0} f_n(d) \geq \sup_{d \geq 0} f(d)$. First we find a random variable $T_{K,n}$ such that

$$\sup_{d \geq 0} f_n(d) = \max_{d \in [0, M_n]} f_n(d),$$

where $M_n = \frac{1}{m} T_{K,n}$ and $m > 0$ is chosen so that $d \geq n$ for all s and n. Hence

$$f_n(d) \leq nK - K \log(1 + dm) + W_q.$$

Denote by

$$T_{K,n} = \frac{1}{m} \left(\exp \left(\frac{n}{n - p - K} F_{K,n} \right) - 1 \right),$$

and observe that for $d > T_{K,n}$ we have $f_n(d) < W_q$ which shows that $T_{K,n}$ has the desired property. Observe now that for every $M > 0$ and for every $t > 0$

$$P \left(\sup_{d \geq 0} f_n(d) \leq t \right) \leq P \left(\max_{d \in [0, M]} f_n(d) \leq t \right).$$

Taking lim sup for $n \to \infty$ one obtains

$$\limsup_{n \to \infty} P \left(\sup_{d \geq 0} f_n(d) \leq t \right) \leq \limsup_{n \to \infty} P \left(\max_{d \in [0, M]} f_n(d) \leq t \right).$$

Because $f_n(d)$ on $C[0, M]$ and by extension on $C[0, \infty)$ one can apply the Continuous Mapping Theorem for the right hand side of the inequality and we obtain

$$\lim_{n \to \infty} P \left(\max_{d \in [0, M]} f_n(d) \leq t \right) = P \left(\max_{d \in [0, M]} f(d) \leq t \right).$$

Using that for any two events A and B, $P(A \cap B) = P(A) - P(A^c \cap B)$, we obtain the following sequence of relations:

$$P \left(\sup_{d > 0} f_n(d) \leq t \right) \geq P \left(\sup_{d > 0} f_n(d) \leq t, T_{K,n} < M \right) = P \left(\max_{d \in [0, M]} f_n(d) \leq t, T_{K,n} < M \right) \geq P \left(\max_{d \in [0, M]} f_n(d) \leq t \right) - P(T_{K,n}$$

Taking the limit when $n \to \infty$ in the first and last expressions we obtain:

$$\liminf_{n \to \infty} P \left(\sup_{d > 0} f_n(d) \leq t \right) \geq P \left(\sup_{d > 0} f(d) \leq t \right) - \lim_{n \to \infty} P(T_{K,n} \geq M),$$

where we used equation (3) and $T_K = \exp \left(\sum_{s=K+1}^{n-p} w_s^2 / K \right) - 1/M$. Consider now a sequence of integers $M \to \infty$. Then $\lim_{M \to \infty} P(T_K \geq M) = 0$. Therefore if we prove that:

$$\lim_{M \to \infty} P \left(\max_{d \in [0, M]} f_n(d) \leq t \right) = P \left(\sup_{d > 0} f(d) \leq t \right),$$

9

then it follows that $\lim_{n \to \infty} P\left(\sup_{d>0} f_n(d) \le t\right)$ exists and:

$$\lim_{n \to \infty} P\left(\sup_{d>0} f_n(d) \le t\right) = P\left(\sup_{d>0} f(d) \le t\right),$$

proving that:

$$\sup_{d>0} f_n(d) = \sup_{d>0} f(d).$$

Denote by $A_M = \{\max_{d \in [0,M]} f(d) \le t\}$. Then $A_M \supset A_{M+1}$ and $\lim_{M \to \infty} P(A_M) = P\left(\sup_{d>0} f(d) \le t\right)$. But $\bigcap_{M=1}^{\infty} A_M = \{\sup_{d>0} f(d) \le t\}$ which ends the proof of equation (4). Observe now that:

$$LRT_n = \sup_{d>0} LRT_n(d) = \sup_{d>0} f_n(d) + V_n.$$

Because V_n converges almost surely to 0 it follows that:

$$LRT_n = \sup_{d>0} f_n(d) = \sup_{d>0} LRT_n(d) + W_q.$$

2.4 Conclusion

3 Proofs of Lemma 3, Theorem 1, and Corollary 1

Before we begin, in this scenario, our null hypothesis is $\lambda = 0$ where λ is the effect size of familial relatedness (FR) and the alternative is $\lambda > 0$.

3.1 Lemma 3

Suppose that $\sup_{s \in \Omega_0}\{\rho_s\} \le B$ are bounded. Under the null hypothesis, (i) both $F_n(\lambda)$ and $G_n(\lambda)$ uniformly converge in probability to $F_0(\lambda) = \sum_{j=0}^{m} q_j \log(w_j) - \log(\sum_{j=0}^{m} q_j w_j$ over $\lambda \in [-\delta, T]$ for any $0 < \delta < \frac{1}{B \vee max_j \{\phi_j\}}$ and $0 < T < +\infty$; (ii) $F_0(\lambda)$ achieves its unique maximum at $\lambda = 0$; (iii) $\hat{\lambda}_n \to^p 0$ and $\hat{\lambda}_n^r \to^p 0$.

  i) Essentially we will be showing that:

$$F_n(\lambda) \to^p F_0(\lambda)$$
$$G_n(\lambda) \to^p F_0(\lambda)$$

where:

$$F_0(\lambda) = \sum_{j=0}^{m} \log(w_j) - \log\left(\sum_{j=0}^{m} q_j w_j\right)$$

10

Additionally, note that ϕ_j is the set of all distictive non-zero eigenvalues of the FR correlation matrix.

Let's start with rewriting $F_n(\lambda)$:

$$F_n(\lambda) = \log\left(\frac{\sum_{i=1}^{n-p} u_i^2}{\sum_{j=0}^{m} f_j' w_j U_j + \sum_{s\in\Omega_0} v_s u_s^2}\right) + \frac{1}{n}\sum_{j=1}^{m} f_j \log(w_j)$$

$$= \log\left(\frac{n-p}{n-p}\cdot\frac{\sum_{i=1}^{n-p} u_i^2}{\sum_{j=0}^{m} f_j' w_j U_j + \sum_{s\in\omega_0} v_s u_s^2}\right) + \frac{1}{n}\sum_{j=1}^{m} f_j \log(w_j)$$

$$= \log\left(\frac{1}{n-p}\sum_{i=1}^{n-p} u_i^2\right) - \log\left(\frac{\sum_{j=0}^{m} f_j' w_j U_j}{n-p} + \frac{\sum_{s\in\Omega_0} v_s u_s^2}{n-p}\right) + \frac{1}{n}\sum_{j=1}^{m} f_j \log(w_j)$$

Looking at the first term, let's note that $u_i \sim N(0,1)$ for $i = 1, ..., n-p$. This allows us to rewrite the first term and note that:

$$\frac{1}{n-p}\sum_{i=1}^{n-p} u_i^2 = \frac{1}{n-p}\sum_{i=1}^{n-p}(u_i - 0)^2 = S_u \to^p \sigma_u = 1$$

$$\frac{1}{n-p}\sum_{i=1}^{n-p} u_i^2 \to^p 1$$

by the week law of large numbers.

$$\Rightarrow \log\left(\frac{1}{n-p}\sum_{i=1}^{n-p} u_i^2\right) \to^p \log(1) = 0$$

Looking at the second term, we note that $U_j = \frac{1}{f_j'}\sum_{i=\Omega_j} u_i^2$ for $j = 0, ..., m$. Additionally, we note here that $f_j' = |\Omega_j|$ which is the count of the times that phi_j is replicated in non-zero eignevalues. For the purposes of this, it is sufficient to say that $|\Omega_j|$ is a count.

We can rewrite U_j as:

$$U_j = \frac{1}{|\Omega_j|}\sum_{i=\Omega_j}(u_i - 0)^2 \to^p 1$$

since $u_i \sim N(0,1)$ as previously stated.

Additionally, this implies that as $U_j \to^p 1$:

$$\Rightarrow \sum_{s\in\Omega_0}\frac{v_s u_s^2}{n-p} \to^p 0$$

under the null hypothesis.

This is since $v_s = 1$ under the null hypothesis:

$$v_s = \frac{1}{1 + \lambda\rho_s} = \frac{1}{1+0} = 1$$

11

Additionally, as long as $\sum_{s \in \Omega_0} u_s^2$ is finite, then $\sum_{s \in \Omega_0} \frac{v_s u_s^2}{n-p} = \sum_{s \in \Omega_0} \frac{u_s^2}{n-p} \to 0$ as $(n-p) \to \infty$.

Noted in the paper:

$$\frac{f_j'}{n-p} \to q_j$$
$$\frac{f_j}{n} \to q_j$$

The suggests that the second term of the equation:

$$-\log\left(\frac{\sum_{j=0}^m f_j' w_j U_j}{n-p} + \frac{\sum_{s \in \Omega_0} v_s u_s^2}{n-p}\right) \to^p -\log\left(\sum_{j=0}^m q_j w_j + 0\right) = -\log\left(\sum_{j=0}^m q_j w_j\right)$$

The third element of the equation converges:

$$\sum_{j=1}^m \frac{f_j}{n} \log(w_j) \to^p \sum_{j=1}^m q_j \log(w_j)$$

Let's now look $G_n(\lambda)$. Let's start by rewriting $G_n(\lambda)$:

$$G_n(\lambda) = \log\left(\frac{\sum_{i=1}^{n-p} u_i^2}{\sum_{j=0}^m f_j' w_j U_j + \sum_{s \in \Omega_0} v_s u_s^2}\right) + \frac{1}{n-p}\sum_{j=1}^m f_j' \log(w_j) + \frac{1}{n-p}\sum_{s \in \Omega_0} \log(v_s)$$

$$= \log\left(\frac{n-p}{n-p} \cdot \frac{\sum_{i=1}^{n-p} u_i^2}{\sum_{j=0}^m f_j' w_j U_j + \sum_{s \in \Omega_0} v_s u_s^2}\right) + \frac{1}{n-p}\sum_{j=1}^m f_j' \log(w_j) + \frac{1}{n-p}\sum_{s \in \Omega_0} \log(v_s)$$

$$= \log\left(\frac{1}{n-p}\sum_{j=1}^{n-p} u_i^2\right) - \log\left(\frac{1}{n-p}\left(\sum_{j=0}^m f_j' w_j U_j + \sum_{s \in \Omega_0} v_s u_s^2\right)\right) + \sum_{j=1}^m \frac{f_j'}{n-p} \log(w_j) + \frac{\sum_{s \in \Omega_0} \log(v_s)}{n-p}$$

In the previous proof of $F_n(\lambda)$ we showed the convergence of the first and second elements of $G_n(\lambda)$.

For the third element, recall that $\frac{f_j'}{n-p} \to q_j$ which implies that $\sum_{j=1}^m \frac{f_j'}{n-p} \log(w_j) \to^p \sum_{j=1}^m q_j \log(w_j)$ which has been shown that the supremum of this third element converges in probability to 0.

The final element of this equation converges to 0, since under the null hypothesis $v_s \to 1$ as previously shown.

Therefore by the uniform convergence theorem, $\sup_{\lambda \in [-\delta, T]} |G_n(\lambda) - F_0(\lambda)| \to^p 0$. $G_n(\lambda)$ therefore converges uniformly to $F_0(\lambda)$ in probability over $\lambda \in [-\delta, T]$.

ii) Recall that $F_0(\lambda) = \sum_{j=0}^m q_j \log(w_j) - \log\left(\sum_{j=0}^m q_j w_j\right)$.

The proof that the authors used involved Jensen's inequality to show that $F_0(\lambda)$ achieves its unique maximum at $\lambda = 0$. I was confused by this and will include my thought process in coming to the conclusion that I did.

When evaluating under the null hypothesis at $\lambda = 0$, we really only need to worry about the second term $-\log(\sum_{j=0}^m q_j w_j)$. Additionally, when considering with the constrain of the bounded parameter

12

space $\lambda \geq 0$, we can see that $F_0(\lambda) \leq 0$ when $\lambda \geq 0$. We also note that q_j is a proportion such that $\sum_{j=0}^{m} q_j \leq 1$.

Under these constraints.

$$F_0(0) = -\log \left(\sum_{j=0}^{m} q_j \right) \mid \sum_{j=0}^{m} q_j \leq 1$$

$$\Rightarrow F_0(0) \geq 0 \ when \ \lambda = 0$$

Since λ can only be non-negative with the way it has been defined in this paper:

$$\left| \sum_{j=0}^{m} q_j \log \left(\frac{1}{1 + \lambda \phi_j} \right) \right| > \left| \log \left(\sum_{j=0}^{m} q_j \frac{1}{1 + \lambda \phi_j} \right) \right|$$

$$\Rightarrow F_0(\lambda) \leq 0$$

Which suggests that $\lambda = 0$ is the unique maximum of $F_0(\lambda)$.

iii) Since $\lambda = 0$ has been established as the unique global maximum, we can reference theorem 5.7 of *Asymptotic Statistics* (1998) by A. van der Vaart, which states that if there is a unique global estimator with the properties we have already established, then $\hat{\theta} \to^p \theta_0$. In this case, both the MLE $\hat{\lambda}_n$ and the REML $\hat{\lambda}_n^r$ have the unique global maximu at $\lambda = 0$ and would both converge in probability to 0 according to this theorem.

3.2 Theorem 1

Suppose the $\sup_{s \in \Omega_0} \leq B$ ***are bounded and*** $p = o\left(\sqrt{n-p}\right)$***. Then, under the null hypothesis (i)*** $\sqrt{n-p}\hat{\lambda}_n \to^p \frac{1}{2}1_{\{0\}} + \frac{1}{2}N(0, \sigma_\lambda^2)^+$***, as*** $n \to \infty$***; (ii)*** $\sqrt{n-p}\hat{\lambda}_n^r \to^D \frac{1}{2}1_{\{0\}} + \frac{1}{2}N(0, \sigma_\lambda^2)^+$***, as*** $(n-p) \to \infty$***; where*** $\sigma_\lambda^2 = \frac{2}{V(Z_\phi)}$ ***and*** $N(0, \sigma_\lambda^2)^+ = max(0, N(0, \sigma_\lambda^2))$ ***is the left truncated distribution of*** $N(0, \sigma_\lambda^2)$ ***at 0.***

As previoiusly discussed in Lemma 3, $\hat{\lambda}_n$ and $\hat{\lambda}_n^r$ are the global maximizers under the null over the interval $[-\delta, T]$.

i)

Let's think of MLE $\hat{\lambda}_n$ as the left truncated at 0. Such that:

$$\hat{\lambda}_n = \{ \begin{array}{ll} \tilde{\lambda}_n & \tilde{\lambda}_n > 0 \\ 0 & \tilde{\lambda}_n \leq 0 \end{array}$$

When n is large enough, note that $F_n'(\tilde{\lambda}_n) = 0$.

If we take the Taylor series expansion of F_n' around 0 we get:

$$0 = F_n'(0) + \tilde{\lambda}_n F_n''(0) + \frac{1}{2}\tilde{\lambda}_n^2 F_n''(c_n \tilde{\lambda}_n)$$

13

where $c \in [0,1]$ and we define $c_n \tilde{\lambda}_n = \eta_n$

We can rewrite the Taylor expansion.

$$-F_n'(0) = \tilde{\lambda}_n F_n''(0) + \frac{1}{2} \tilde{\lambda}_n^2 F_n'''(\eta_n)$$

$$\Rightarrow \tilde{\lambda}_n \left(F_n''(0) + \frac{1}{2} \tilde{\lambda} F_n'''(\eta_n) \right) = -F_n'(0)$$

$$\Rightarrow \sqrt{n-p}\, \tilde{\lambda}_n = \frac{\sqrt{n-p}\, F_n'(0)}{-F_n''(0) - \frac{1}{2} \tilde{\lambda}_n F_n'''(\eta_n)}$$

We can find $F_n'(0)$ and $F_n''(0)$ by taking the derivatives with respect to λ and evaluating at $\lambda = 0$.

$$F_n'(0) = \frac{\partial}{\partial \lambda} \left[\frac{\sum_{i=1}^{n-p} u^2}{\sum_{j=0}^{m} f_j \frac{1}{1+\lambda \phi_j} U_j + \sum_{s \in \Omega_0} \frac{1}{1+\lambda \rho_s} u_s^2} + \frac{1}{n} \sum_{j=1}^{m} f_j \log \left(\frac{1}{1+\lambda \phi_j} \right) \right]_{\lambda=0}$$

$$= \frac{\sum_{j=0}^{m} \frac{f_j}{n-p} \phi_j U_j + \sum_{s \in \Omega_0} \frac{\rho_s u_s^2}{n-p}}{\sum_{j=0}^{m} \frac{f_j'}{n-p} U_j + \sum_{s \in \Omega_0} \frac{u_s^2}{n-p}} - \sum_{j=1}^{m} \frac{f_j}{n} \phi_j$$

$$F_n''(0) = \left(\frac{\sum_{j=0}^{m} \frac{f_j}{n-p} \phi_j U_j + \sum_{s \in \Omega_0} \frac{\rho_s u_s^2}{n-p}}{\sum_{j=0}^{m} \frac{f_j'}{n-p} U_j + \sum_{s \in \Omega_0} \frac{u_s^2}{n-p}} \right)^2 - \frac{2 \sum_{j=0}^{m} \frac{f_j'}{n-p} \phi_j^2 U_j + 2 \sum_{s \in \Omega_0} \frac{\rho_s^2 u_s^2}{n-p}}{\sum_{j=0}^{m} \frac{f_j'}{n-p} U_j + \sum_{s \in \Omega_0} \frac{u_s^2}{n-p}} + \sum_{j=1}^{m} \frac{f_j}{n} \phi_j^2$$

Let $\bar{\rho}_n = \sum_{j=0}^{m} \frac{f_j'}{n-p} \phi_j$ and $\bar{\zeta}_n = \sum_{j=0}^{m} \frac{f_j}{n} \phi_j$.

We can now find what the different parts of the equation converge to.

$$\sqrt{n-p}\, F_n'(0) = \sqrt{n-p} \left(\frac{\sum_{j=0}^{m} \frac{f_j'}{n-p} \phi_j U_j + O\left(\frac{p}{n-p} \right)}{\sum_{j=0}^{m} \frac{f_j'}{n-p} U_j + O\left(\frac{p}{n-p} \right)} - \bar{\zeta}_n \right)$$

$$= \frac{\sum_{j=0}^{m} \sqrt{\frac{f_j'}{n-p}} (\phi_j - \bar{\zeta}_n) \sqrt{f_j'}(U_j - 1) + \sqrt{n-p}(\bar{\rho}_n - \bar{\zeta}_n) + O\left(\frac{p}{\sqrt{n-p}} \right)}{\sum_{j=0}^{m} \frac{f_j'}{n-p} U_j + O\left(\frac{p}{n-p} \right)}$$

Here we note that $\bar{\rho}_n \to \sum_{j=0}^{m} q_j \phi_j = \bar{\phi}$ and $\bar{\zeta}_n \to \bar{\phi}$.

Additionally, we note the following:

$$\sqrt{f_j'}(U_j - 1) \to^D z_i = N(0,2)$$

$$\sum_{j=0}^{m} \frac{f_j'}{n-p} U_j + O\left(\frac{p}{n-p} \right) \to^p 1$$

Also, let $u^* = \lim_{n \to +\infty} \sqrt{n-p}(\bar{\rho}_n - \bar{\zeta}_n) = \lim_{n \to +\infty} u_n = 0$. Recall that in the paper $V(Z_\phi) = \sum_{j=0}^{m} q_j (\phi_j - 1)^2$

14

Under those conditions and by Slutsky's theorem

$$\sqrt{n-p}F_n'(0) \to^D N(u^*, 2V(Z_\phi))$$

For the second part of the equation, note that:

$$F_n''(0) \to^P F_n''(\lambda_0) = -V(Z_\phi)$$

We note that $F_n'''(\eta_n)$ is finite and the $\hat{\lambda}_n \to^P 0$ as defined by Lemma 3. This implies that $\hat{\lambda}_n F_n'''(\eta_n) \to^P 0$.

This lets us say that

$$\sqrt{n-p}\tilde{\lambda}_n \to^D N\left(0, \frac{2}{V(Z_\phi)}\right)$$

And since this hypothesis test is bounded on the edge of the parameter space, we can say that:

$$\sqrt{n-p}\hat{\lambda}_n \to^D \frac{1}{2}(1) + \frac{1}{2}N\left(0, \frac{2}{V(Z_\phi)}\right)$$

ii)

This proof is pretty similar to part i), with a few minor changes since we will be working with $G_n(\lambda)$. We essentially follow the same steps.

When doing the Taylor's expansion around 0 and solving for $\tilde{\lambda}_n^r$ we get:

$$\sqrt{n-p}\tilde{\lambda}_n^r = \frac{\sqrt{n-p}G_n'(0)}{-G_n''(0) - \frac{1}{2}\tilde{\lambda}_n^r G_n''(\eta_n)} - \frac{\sum_{s\in\Omega_0}\rho_s v_s^2}{n-p}$$

Taking the derivatives of $G_n(\lambda)$ and evaluating at 0 to find $G_n'(0)$ and $G_n''(0)$:

$$G_n'(0) = \frac{\sum_{j=0}^m \frac{f_j'}{n-p}\phi_j U_j + \frac{1}{n-p}\sum_{s\in\Omega_0}\rho_s u_s^2}{\sum_{j=0}^m \frac{f_j'}{n-p}U_j + \frac{1}{n-p}\sum_{s\in\Omega_0}u_s^2} - \sum_{j=1}^m \frac{f_j'}{n-p}\phi_j - \sum_{s\in\Omega_0}\frac{\rho_s}{n-p}$$

$$G_n''(0) = \left(\frac{\sum_{j=0}^m \frac{f_j'}{n-p}\phi_j U_j + \frac{1}{n-p}\sum_{s\in\Omega_0}\rho_s u_s^2}{\sum_{j=0}^m \frac{f_j'}{n-p}U_j + \frac{1}{n-p}\sum_{s\in\Omega_0}u_s^2}\right)^2 - \frac{2\sum_{j=0}^m \frac{f_j'}{n-p}\phi_j U_j + \frac{1}{n-p}\sum_{s\in\Omega_0}\rho_s^2 u_s^2}{\sum_{j=0}^m \frac{f_j'}{n-p}U_j + \frac{1}{n-p}\sum_{s\in\Omega_0}u_s^2} + \sum_{j=1}^m \frac{f_j'}{n-p}\phi_j^2 + \sum_{s\in\Omega_0}\frac{\rho_s^2}{n-p}$$

Under Lemma 1 of the paper, $\sum_{s\in\Omega_0}\rho_s = O(p)$. Also define $\bar{\phi} = \sum_{j=0}^m q_j\phi_j$ and $\bar{\rho}_n = \sum_{j=0}^m \frac{f_j'}{n-p}\phi_j$.

This implies that:

$$\sqrt{n-p}G_n'(0) = \frac{sum_{j=0}^m \sqrt{\frac{f_j'}{n-p}}(\phi_j - \bar{\rho}_n)\sqrt{f_j'}(U_j - 1) + O\left(\frac{1}{\sqrt{n-p}}\right)}{\sum_{j=0}^m \frac{f_j'}{n-p}U_j + O\left(\frac{p}{n-p}\right)} + O\left(\frac{p}{\sqrt{n-p}}\right)$$

We note the following:

15

$$\sqrt{f'_j}(U_j - 1) \to^D z_j \sim N(0, 2)$$

$$\sum_{j=0}^m \frac{f'_j}{n-p} U_j + O\left(\frac{p}{n-p}\right) \to^p 1$$

$$\Rightarrow \sqrt{n-p} G'_{n}(0) \to^D \sum_{j=0}^m \sqrt{q_j}(\phi_j - \bar{\phi}) z_j$$

Additionally note that:

$$G''_n(0) \to^p \phi^2 - \sum_{j=1}^m q_j \phi_j^2 = -V(Z_\phi)$$

And, since $G'''_n(\eta_n)$ is finite we can say that $\tilde{\lambda}_n^r G'''_n(\eta_n) \to^p 0$

By Slutsky's:

$$\sqrt{n-p}\tilde{\lambda}_n^r \to^D N(0, \frac{2}{V(Z_\phi)})$$

$$\sqrt{n-p}\hat{\lambda}_n^r \to^D \frac{1}{2} 1 + \frac{1}{2} N(0, \frac{2}{V(Z_\phi)})$$

3.3   Corollary 1

Under the same assumptions as in Theorem 1, i) $LRT_n \to^D \frac{1}{2} 1_0 + \frac{1}{2} \chi_1^2$, as $n \to \infty$; ii) $RLRT_n \to^D \frac{1}{2} 1_{\{0\}} + \frac{1}{2} \chi_1^2$, as $(n-p) \to \infty$.

i)

If we take a Taylor Expansion around $F_n(\tilde{\lambda}_n)$ gives us:

$$F_n(\tilde{\lambda}_n) = \tilde{\lambda}_n(F'_n(0) + \frac{1}{2}\tilde{\lambda}_n F''_n(0) + \frac{1}{6}\tilde{\lambda}_n^2 F'''_n(\tilde{\eta}_n))$$

where $\tilde{\eta}_n = \tilde{c}_n \tilde{\lambda}_n$ and $\tilde{c}_n \in [0, 1]$.

Recall that according to the proof for Theorem 1, $F'_n(0)$ can be written as:

$$F'_n(0) = -\tilde{\lambda}_n F''_n(0) - \frac{1}{2}\tilde{\lambda}_n^2 F'''_n(\eta_n)$$

We can combine the two equations and get:

$$(n-p)F_n(\tilde{\lambda}_n) = -\frac{1}{2}(n-p)\tilde{\lambda}_n^2 F''_n(0) + (n-p)\tilde{\lambda}_n^3 \left[\frac{1}{6}F'''_n(\tilde{\eta}_n) - \frac{1}{2}F'''_n(\eta_n)\right]$$

We note here that $F'''_n(\eta_n)$ and $F'''_n(\tilde{\eta}_n)$ are both finite and $\tilde{\lambda}_n \to^p 0$ which allows us to rewrite the equation as:

$$(n - p)F_n(\tilde{\lambda}_n) = -\frac{F_n''(0)}{2}(\sqrt{n-p}\tilde{\lambda}_n)^2$$

Recall from the previous proof that $-F_n''(0) \to^p V(Z_\phi)$ and $\sqrt{n-p}\tilde{\lambda} \to^D N\left(0, \frac{2}{V(Z_\phi)}\right)$.

Thus, when $\tilde{\lambda}_n > 0$, $(n-p)F_n(\tilde{\lambda}_n) \to^D \chi_1^2$ and

$$LRT_n \to^D \frac{1}{2}1_{\{0\}} + \frac{1}{2}\chi_1^2$$

as $n \to \infty$.

ii)

For the RLRT, we start with doing a Taylor's expansion of G_n around 0:

$$G_n(\tilde{\lambda}_n^r) = \tilde{\lambda}_n^r(G_n'(0) + \frac{1}{2}\tilde{\lambda}_n^r G_n''(0) + \frac{1}{6}(\tilde{\lambda}_n^r)G_n'''(\tilde{\eta}_n))$$

Where $\tilde{\eta}_n = \tilde{c}_n\tilde{\lambda}_n$ and $\tilde{c}_n \in [0, 1]$.

Note that from the Theorem 1 proof we can write $G - n'(0)$ as:

$$G_n'(0) = -\tilde{\lambda}_n^r G_n''(0) - \frac{1}{2}(\tilde{\lambda}_n^r)^2 G_n'''(\eta_n)$$

We can use this equation to get:

$$(n-p)G_n(\tilde{\lambda}_n^r) = -\frac{1}{2}(n-p)(\tilde{\lambda}_n^r)^2 G_n''(0) + (n-p)(\tilde{\lambda}_n^r)^3 \left[\frac{1}{6}G_n'''(\tilde{\eta}_n) - \frac{1}{2}G_n'''(\eta_n)\right]$$

Since $G_n'''(\tilde{\eta}_n)$ and $G_n'''(\eta)$ are finite and $(\tilde{\lambda}_n^r) \to^p 0$ we can rewrite this as:

$$(n-p)G_n(\tilde{\lambda}_n^r) = -\frac{1}{2}(n-p)(\tilde{\lambda}_n^r)^2 G_n''(0)$$

Which can again be rewritten as:

$$(n-p)G_n(\tilde{\lambda}_n^r) = \left(-\frac{G_n''(0)}{V(Z_\phi)}\right)\left(\frac{V(Z_\phi)(n-p)}{2}(\tilde{\lambda}_n^r)^2\right)$$

From the proof on Theorem 1, we can note that $-G_n''(0) \to^p V(Z_\phi)$. The second term in the equation, $\left(\frac{V(Z_\phi)(n-p)}{2}(\tilde{\lambda}_n^r)^2\right) \to \chi_1^2$.

Therefore, when $\tilde{\lambda}_n^r > 0$:

$$(n-p)G_n(\tilde{\lambda}_n^r) \to^D \chi_1^2$$
$$\Rightarrow RLRT \to^D \frac{1}{2}1_{\{0\}} + \frac{1}{2}\chi_1^2$$

4 Simulation Study

4.1 Balanced one-way analysis of variance

4.1.1 Algorithm to simulate the null finite sample distribution of LRT_n

Step 1: define a grid $0 = \lambda_1 < \lambda_2 < \ldots < \lambda_m$ of possible values for λ.

Step 2: simulate K independent χ_1^2 random variables $w_1^2, \ldots, w_K^2$. Set $S_K = \sum_{s=1}^{K} w_s^2$.

Step 3: independently of step 1, simulate $X_{n,K,p} = \sum_{s=K+1}^{n-p} w_s^2$ with a χ_{n-p-K}^2-distribution.

Step 4: independently of steps 1 and 2, simulate $X_q = \sum_{s=1}^{q} u_s^2$ with a χ_q^2-distribution.

Step 5: for every grid point λ_i compute

$$N_n(\lambda_i) = \sum_{s=1}^{K} \frac{\lambda_i \mu_{s,n}}{1 + \lambda_i \mu_{s,n}} w_s^2$$

$$D_n(\lambda_i) = \sum_{s=1}^{K} \frac{w_s^2}{1 + \lambda_i \mu_{s,n}} w_s^2 + X_{n,K,p}$$

Step 6: determine λ_{max} which maximizes $f_n(\lambda_i)$ over $\lambda_1, \ldots, \lambda_m$.

Step 7: compute

$$LRT_n = f_n(\lambda_{max}) + n \log(1 + \frac{X_q}{S_K + X_{n,K,p}})$$

Step 8: repeat steps 2-7 until the desired number of simulations is achieved.

5 Conclusion