Stat 455 Lecture 13: Model Check and Diagnostics in Linear Mixed Models (10-09-2007)

1 Goodness-of-fit and Best Model

1) An unnecessarily detailed model (beyond a correct model) may still have good fit (goodness-of-fit is high).

2) However, we prefer the correct model which is also economical in parameters.

2 Pearson's χ^2-test for $H_0 : F = F_0$

We understand the following formula:

$$\hat{\chi}_P^2 = \sum_{k=1}^{M} \frac{(N_k - EN_k)^2}{EN_k} \tag{1}$$

Remark When EN_k is to be calculated from estimated model parameters, the asymptotic null-distribution of $\hat{\chi}_P^2$ may not be χ^2 (Moore, 1978).

3 Hypothesis Test (Jiang, 2001)

3.1 Null Hypothesis

For model

$$Y = X\beta + Z_1\alpha_1 + \cdots + Z_s\alpha_s + \epsilon, \tag{2}$$

notations are given in Figure 1. We are testing the null hypothesis in Figure 2.

3.2 Proposed Test Statistic

$$\hat{\chi}^2 = \frac{1}{a_n} \sum_{k=1}^{M} (N_k - E_{\hat{\theta}} N_k)^2 \tag{3}$$

where,

1) a_n is a normalizing constant.

2) N_k= ? $E_{\hat{\theta}}N_k$=?

3) $\hat{\theta}$ is the estimator for parameters.

4) The observed counts N_k is NOT a sum of independent random variables, while correlated in LMM.

5) Normalizing constant is **same** for all squares for simplicity.

3.3 Results

1) Under regularity conditions, the asymptotic distribution of $\hat{\chi}^2$ is a weighted χ^2, i.e., $\sum_{j=1}^{M} \lambda_j Z_j^2$, where Z_js are independent N(0,1).

2) $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_M$ are eigenvalues of some nonnegative definite matrix $\Sigma_N(\theta)$ (pages 112-113 in textbook, estimated as $\Sigma_N(\hat{\theta})$).

3) Ideally, $c_\rho(\lambda_1, \ldots, \lambda_M)$ is the ρ-critical value of the random variables $\sum_{j=1}^{M} \lambda_j Z_j^2$, where

4) $\Pr(\hat{\chi}^2 > \hat{c}_\rho) \to \rho$ $(n \to \infty)$, where $\hat{c}_\rho = c_\rho(\hat{\lambda_1}, \ldots, \hat{\lambda_M})$.

5) The selection of normalizing constant a_n (Jiang 2001, page 92 in textbook).

4 Model Selection: with Fixed Random Factors

Figure 3 and Figure 4 (Large space).

5 References

◇ Jiang, J. (2001), Goodness-of-fit tests for mixed model diagnostics, Annals of Statistics, 29(4), 1137-1164.

◇ Jiang, J. and Rao, J.S.(2003), Consistent procedures for mixed linear model selection, Sankhya 65, 23-42.

6 Appendix

$$y = (y_i)_{1 \le i \le N}$$

$$X = N \times p \text{ matrix}$$

$$\alpha_r = (\alpha_{rk})_{1 \le k \le m_r}, \ \underset{\sim}{iid} \ F_r(\cdot \mid \sigma_r), \ \boxed{\begin{matrix} \text{Mean} = 0 \\ \text{Var} = \sigma_r^2 \end{matrix}}$$

$$(r = 1, 2, \cdots, s)$$

$$\varepsilon = (\varepsilon_i)_{1 \le i \le N}, \ \underset{\sim}{iid} \ G(\cdot \mid \tau) \ \boxed{\begin{matrix} \text{Mean} = 0 \\ \text{Var} = \tau^2 \end{matrix}}.$$

Figure 1: Notations: model

$$H_0: F_r(\cdot \mid \sigma_r) = F_{0r}(\cdot \mid \sigma_r), \quad 1 \le r \le s$$

and

$$G(\cdot \mid \tau) = G_0(\cdot \mid \tau)$$

$\boxed{\text{remark}}$: σ_r, τ^2, β are unknown.

Figure 2: Hypothesis Test

Criterion:

$$C_n(a) = \left|y - X(a)\hat{\beta}(a)\right|^2 + \lambda_n(a) \qquad (*)$$

$$= \left|P^{\perp}_{X(a)}\, y\right|^2 + \lambda_n |a|$$

$$\begin{cases} a \in \mathcal{A} \text{ (subset of models containing true model)} \\ |a| = \text{cardinality of } a. \end{cases}$$

True Model $\boxed{a_0}$

$$\begin{cases} \nu_n = \max_{1 \le j \le p} |X_j|^2 \\ \beta = (\beta_j)_{1 \le j \le p}, \ (\beta_j \neq 0) \\ \rho_n = \lambda_{max}(ZGZ_1) + \lambda_{max}(R), \end{cases}$$

Then $\boxed{\begin{array}{l} \text{minimizer } \hat{a} \text{ of } (*) \text{ over } a \in \mathcal{A}: \\ Pr(\hat{a} \neq a_0) \to 0 \ (n \to \infty), \text{ if} \\ \lambda_n / \nu_n \to 0 \text{ and } \rho_n / \lambda_n \to 0 \end{array}}$

Figure 3: Model selection (result 1)

When, dimension is $\boxed{\text{Large}}$.

$$X\beta = \sum_{j=1}^{\ell} \beta_j X_j$$ with some β_js may be zero,

True model $\boxed{a_0}$: $\{1 \le j \le \ell;\ \beta_j \ne 0\}$

Let

$$\begin{cases} X_{-j} = (X_u)_{1 \le u \le \ell,\, u \ne j},\ j \in \overline{1,\ell}, \\ \eta_n = \min_{1 \le j \le \ell} \left| P^{\perp}_{X_{-j}}(X_j) \right|^2, \quad \boxed{P^{\perp}_M = I - P_M} \\ \delta_n : \text{positive numbers} \end{cases}$$

Let $\hat{a}$ be the subset of $I = \{1, \cdots, \ell\}$ s.t.

$$\frac{\left(|P^{\perp}_{X_{-j}} y|^2 - |P^{\perp}_{X} y|^2 \right)}{\left(|P^{\perp}_{X_{-j}} X_j|^2 \delta_n \right)} > 1$$

for $j \in \hat{a}$, then

$$\boxed{\begin{array}{l} \text{If } \rho_n/\eta_n \to 0,\ \delta_n \to 0,\ \frac{\rho_n}{\eta_n \delta_n} \to 0 \Rightarrow \\ \quad \hat{a} \text{ is consistent.} \end{array}}$$

Figure 4: Model selection (result 2)