Project 3 of ST 755, Fall 2011

Due: Monday, 11/21/2011

1. Consider the following special linear mixed model:

$$Y_{ij} = \beta + b_i + e_{ij} \qquad i = 1, ..., m, \qquad j = 1, 2,$$

where $b_i \sim \mathrm{N}(0, \sigma_b^2)$ and $e_{ij} \overset{\text{iid}}{\sim} \mathrm{N}(0, \sigma_e^2)$, independent of b_i. Suppose the true values of σ_b^2 and σ_e^2 are $\sigma_b^2 = 0$ and $\sigma_e^2 = 1$ (the true value of β is not very important for this problem). Denote $\theta = (\beta, \sigma_b^2, \sigma_e^2)^T$ and let θ_0 denote its true value. Do the following:

   (a) Show that the Fisher information matrix $I(\theta_0)$ for the MLE of θ at θ_0 is equal to

$$I(\theta_0) = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 1 & 1 \end{bmatrix}.$$

   Note that $I(\theta_0)$ is calculated at the true value θ_0. Here m is the same as the n in Self and Liang (1987), and $Y_i = (Y_{i1}, Y_{i2})^T$ is the same as X_i.

   (b) Show that the asymptotic distribution of the MLE of β will not be affected by the boundary issue. That is, $\sqrt{m}(\widehat{\beta} - \beta_0)$ will still have an asymptotic normal distribution with mean zero. Find the variance.

   (c) However, the asymptotic distribution of the MLEs of σ_b^2 and σ_2^2 will be affected by the boundary issue. That is, $\sqrt{m}(\widehat{\sigma}_b^2 - 0)$ and $\sqrt{m}(\widehat{\sigma}_e^2 - 1)$ will have asymptotic distributions with non-zero means (so the biases of $\widehat{\sigma}_b^2$ and $\widehat{\sigma}_e^2$ are at the order $m^{-1/2}$). Find the means for these distributions. We know $\widehat{\sigma}_b^2$ will over-estimate its true value ($= 0$). Will $\widehat{\sigma}_e^2$ also over-estimate its true value ($= 1$)?

2. Consider the following linear mixed model for longitudinal data

$$Y_{ij} = x_{ij}^T\beta + b_{i0} + b_{i1}t_{ij} + e_{ij} \qquad i = 1, ..., m, \qquad j = 1, 2, ..., n_j,$$

where β are fixed effects (of x), t_{ij} are the times when Y_{ij} are measured, $(b_{i0}, b_{i1})^T$ are random effects independent of the measurement errors e_{ij}. It is further assumed that

$$\begin{pmatrix} b_{i0} \\ b_{i1} \end{pmatrix} \sim \mathrm{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_{00} & \sigma_{01} \\ \sigma_{01} & \sigma_{11} \end{bmatrix} \right)$$

and e_{ij} are iid from $N(0, \sigma_e^2)$. We are interested in conducting the likelihood ratio test for the presence of the random effect b_{i1}.

(a) Specify H_0 and H_1 in terms of model parameters for the above testing problem.

(b) From the results presented in class, we know that under H_0 the LRT statistic for testing the above hypothesis asymptotically has a distribution equal to the 50:50 mixture of χ_1^2 and χ_2^2. Find out the critical values for the LRT for three levels 0.01, 0.05, 0.1;

(c) If we used the regular LRT to test H_0, what is the consequence? Specifically, what are the actual sizes of the test if the nominal levels of the regular LRT are chosen to be 0.01, 0.05, 0.1?

(d) Suppose we observed LRT = 2.45. What is the p-value?

(e) Consider a special case of the above model:

$$Y_{ij} = \beta_0 + \beta_1 t_{ij} + b_{i0} + b_{i1} t_{ij} + e_{ij} \quad i = 1, ..., 100, \quad j = 1, 2, ..., 5,$$

where $t_{ij} = j$. Assume under H_0 the model parameters take values $\beta_0 = 2, \beta_1 = 1, \sigma_{00} = 1, \sigma_e^2 = 2$. Use simulation to get the empirical distributions of the LRT statistic and p-value (use 1000 simulation runs). Compare this empirical distribution with its theoretical distribution. You can use the SAS procedure mixed to conduct the simulation.