Homework #8 -- due Wednesday, 22 April 2009 *** turn in just starred (*) questions ***

Exercises at the end of Chapter 7: 1, 2, 12
Exercises at the end of Chapter 8: 1*, 6*, 7, 8

Two questions from last year's final exam:
*3) For modelling the growth of trees, the following model is often employed:
(*) $y_{it} = \mu + \alpha_i + \beta_i t + e_{ij}$, $i = 1, ..., a$; $t = 1, ...n$,
where y_{it} measures the girth (circumference) of tree i at time t. Here μ represents an average intercept, and α_i, β_i are random intercepts and slopes due to the timing of the emergence of the sprout and the initial tagging of the tree. While the errors would be modelled as usual as e_{it} iid $N(0, \sigma^2)$, here model the random tree effects as

$$\begin{bmatrix} \alpha_i \\ \beta_i \end{bmatrix} \text{ iid } N_2\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix} \right)$$

independently of e_{it}.

a) For the simple case of a = 3 trees, and n = 4 time points, give the appropriate matrices and vectors for **X**, **b**, **Z**, and **u** (four of them), in order to write the model (*) in the general mixed model form:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Zu} + \mathbf{e},$$

where **Xb** represents the fixed effects, and **u** the random effects. (Hint: **u** should be 6×1)

$$\mathbf{y} = (y_{11}, y_{12}, y_{13}, y_{14}, y_{21}, y_{22}, y_{23}, y_{24}, y_{31}, y_{32}, y_{33}, y_{34})^T$$
$$\mathbf{e} = (e_{11}, e_{12}, e_{13}, e_{14}, e_{21}, e_{22}, e_{23}, e_{24}, e_{31}, e_{32}, e_{33}, e_{34})^T$$

b) Give Cov(**u**).
c) Write out the first four rows of Cov(**y**).

*4) In large samples, the observed proportions from a multinomial distribution with sample size n and probabilities p_i, $i = 1, ..., k$ can be approximated by a multivariate normal distribution. Following that motivation, assume that we observe

$$\mathbf{y} = (Y_1, Y_2, ..., Y_k)^T \sim N_k(n\mathbf{P1}_k, n(\mathbf{P} - \mathbf{P11}^T\mathbf{P})),$$

where $\mathbf{P} = \text{diag}(p_1, p_2, ..., p_k)$ so that $\mathbf{P1}_k = (p_1, p_2, ..., p_k)^T$ and also $\sum_{i=1}^{k} p_i = \mathbf{1}^T\mathbf{P1} = 1$.

a) Show that the covariance matrix is singular by finding a vector **v** such that $\text{Var}(\mathbf{v}^T\mathbf{y}) = 0$.
Consider partitioning the observed vector **y** and matrix **P** as follows:

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}_* \\ y_k \end{bmatrix} \begin{matrix} \text{k-1} \\ 1 \end{matrix} \qquad \mathbf{P} = \begin{bmatrix} \mathbf{P}_* & \mathbf{0} \\ \mathbf{0} & p_k \end{bmatrix}$$

so that $\mathbf{y}_*$ has a nonsingular normal distribution.
b) Find the distribution of $\mathbf{1}_{k-1}^T\mathbf{y}_*$.
c) Find a vector $\boldsymbol{\mu}$ and $(k-1) \times (k-1)$ matrix **A** so that $(\mathbf{y}_* - \boldsymbol{\mu})^T\mathbf{A}(\mathbf{y}_* - \boldsymbol{\mu})$ that has a central chi-square distribution. Give its degrees of freedom and a simplified expression for it. (Hints: 1) Use Result 6, and 2) you might simplify this to something very familiar)