STA 2201s06 Assignment 3

Do this assignment in preparation for Quiz Three on Thursday Feb. 9th. The handwritten parts are preparation for the quiz, and are not to be handed in. Question 2 asks for calculations with R. Please bring your printouts to the quiz; they may be handed in.

1. Let $\mathbf{X}_1, \ldots, \mathbf{X}_n$ be a random sample from a multivariate normal population with mean $\boldsymbol{\mu}$ and variance-covariance matrix $\boldsymbol{\Sigma}$. Using the MLEs

$$\widehat{\boldsymbol{\mu}} = \overline{\mathbf{X}} \text{ and } \widehat{\boldsymbol{\Sigma}} = \frac{1}{n}\sum_{i=1}^{n}(\mathbf{X}_i - \overline{\mathbf{X}})(\mathbf{X}_i - \overline{\mathbf{X}})',$$

   derive the large-sample likelihood ratio test G for testing whether the components of the random vectors $\mathbf{X}_i$ are independent. That is, we want to test whether $\boldsymbol{\Sigma}$ is diagonal. It is okay to use material from the class notes without proof.

2. Using R, write a program to compute the test you derived in the preceding question. Your program should return 3 values: G, the degrees of freedom, and the p-value. Run it on the sample in `fourvars.dat`; see link to the data on the course web page. Bring a printout listing your program and illustrating the run on `fourvars.dat`. Of course it would be nice if your program were general, but it is not required. Note that for this problem, numerical maximum likelihood is not needed. Both your restricted and your unrestricted MLEs can and should be in explicit form.

3. Please read Chapter One in the text. Note the author's use of lower case quantities for random variables. This is old-fashioned, and we will not follow this practice. Don't worry too much about the historical part; it is full of references to stuff you have not learned yet. Your task: Derive expression 1.4; show your work.

4. Please read Pages 10-20 in Chapter Two. You will find it a really mixed bag. One moment, he is clearing up misconceptions for some of the older social scientists in the audience, and the next moment he is introducing the notation of the LISREL model, which definitely takes some getting used to even for a statistician. Then he tells you how to pronounce the Greek letters, and explains that a covariance matrix is an "unstandardized correlation matrix." (Page 15). Yikes.

   Well, some of the funny vocabulary is important, especially the following: endogenous, exogenous, latent, observed. Take a look at the notation but do not be concerned; I will cover it thorougly in class. There are no problems on this yet, and nothing from Chapter Two will be on the Quiz.