Likelihood Theory and Extensions (BIOS:7110)
Breheny

Assignment 9
Due: Monday, November 1

1. *Slutsky's extension.* Suppose $\mathbf{y}_n \xrightarrow{\text{d}} \mathbf{y}$, where $\mathbf{y}$ is a $d \times 1$ random vector, $\mathbf{A}_n \xrightarrow{\text{P}} \mathbf{A}$, where $\mathbf{A}$ is a positive definite matrix, and that $\mathbf{y}_n = \mathbf{A}_n \mathbf{x}_n$.

   (a) Prove that $\mathbf{x}_n$ is bounded in probability. Hint: If $\mathbf{A}$ is positive definite, $\mathbf{A}^2$ is positive definite.

   (b) Prove that $\mathbf{x}_n \xrightarrow{\text{d}} \mathbf{A}^{-1}\mathbf{y}$.

2. *Bernstein-von Mises phenomenon.* Consider the posterior distribution of a parameter $\boldsymbol{\theta}$ in a Bayesian model. Suppose regularity conditions (A)-(C) as defined in the "Likelihood: Consistency" lecture are satisfied and that condition (C) also applies to the prior $p(\boldsymbol{\theta})$; i.e., the prior is continuous and smooth in a neighborhood of $\boldsymbol{\theta}^*$. Take a second-order Taylor series expansion of log-posterior about $\hat{\boldsymbol{\theta}}$, the point that maximizes the posterior (for the sake of this problem, you may assume that the posterior is unimodal and that this mode is consistent) to show that the posterior density is approximately that of a multivariate normal distribution with mean $\hat{\boldsymbol{\theta}}$ and variance $(n\boldsymbol{\mathscr{I}}(\boldsymbol{\theta}^*))^{-1}$.

   Specifically, I am asking you to (i) take a Taylor series expansion of $\log p(\boldsymbol{\theta}|\mathbf{x})$ (ii) divide by n and take the limit as $n \to \infty$ (iii) relate the resulting limit back to the density of a multivariate normal distribution.

   Notes: This is a simplified version of the Bernstein-von Mises theorem, in two respects.

   - You are not asked to show consistency of $\hat{\boldsymbol{\theta}}$.
   - You do not have to consider the remainder term of the Taylor series expansion (the third-order term). This remainder becomes rather complicated for the Bayesian posterior; to deal with it properly, one has to consider the posterior distribution of $\sqrt{n}(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}})$ or $\sqrt{n}(\boldsymbol{\theta} - \boldsymbol{\theta}^*)$, both of which converge to random distributions as $n \to \infty$. The underlying cause of this complication is that the posterior distribution (unlike the sampling distribution in frequentist approaches) is a conditional distribution, and therefore random.