UNIVERSITY OF MASSACHUSETTS
Department of Mathematics and Statistics
Advanced Exam - Linear Models
Friday, August 29, 2014

Please work all problems. Seventy points are required to pass.

1. (34 points) State and prove the Gauss-Markov Theorem. Please be sure to define any notation you use, and please state your assumptions.

2. (35 points) Let $\mathbf{y} = (y_1, \ldots, y_n)^{\mathrm{T}}$, let $\mathbf{X}$ be an n by p matrix with rank $q \leq p \leq n, q \geq 1$. Let $\boldsymbol{\beta}$ be a vector of length p. Let $\boldsymbol{\epsilon}$ be a random vector of length n with uncorrelated elements that each have mean zero and the same variance.

   (a) Derive the "normal equations" for $\widehat{\boldsymbol{\beta}}$.

   (b) Suppose that $\boldsymbol{\epsilon}$ is multivariate normal too. Show that a solution to the normal equations for $\widehat{\boldsymbol{\beta}}$ is also a maximum likelihood estimator for $\boldsymbol{\beta}$.

   (c) Suppose that $\boldsymbol{\epsilon}$ is multivariate normal too. Derive a maximum likelihood estimator for σ^2.

   (d) Derive the bias of your estimator from part (c), and propose an unbiased alternative.

   (e) Suppose that $q < p$. Show that the normal equations have more than one solution.

   (f) Give an example of a linear model where $q = 2$ and $p = 3$.

   (g) Give two solutions to the normal equations for your example in part (f).

   (h) Give the definition of an estimable function.

   (i) Please give an example of a maximally sized set of linearly independent estimable functions using your model from part (f).

3. (31 points) For the following set of problems, please use the same notation and assumptions as in (2), but now please assume that $q = p$. Let $\widehat{\boldsymbol{\beta}}$ solve the normal equations. Let $\widehat{\mathbf{y}} = \mathbf{X}\widehat{\boldsymbol{\beta}}$ and $\mathbf{e} = \mathbf{y} - \widehat{\mathbf{y}}$. Suppose that the first column of $\mathbf{X}$ is all 1s.

   (a) Let $\mathbf{1}_n$ be a vector of all 1s with length n. Derive an expression for $\mathbf{1}_n^{\mathrm{T}}\mathbf{e}$.

   (b) Show that $\mathbf{1}_n^{\mathrm{T}}\mathbf{y} = \mathbf{1}_n^{\mathrm{T}}\widehat{\mathbf{y}}$.

   (c) Find the expectations and covariances of $\widehat{\boldsymbol{\beta}}$ and $\widehat{\mathbf{y}}$.

   (d) What is the maximally sized set of linearly independent estimable functions in this case?

   (e) Suppose that $\boldsymbol{\epsilon}$ is multivariate normal as well.

      i. Let $\widehat{\beta}_k$ be the kth element of $\widehat{\boldsymbol{\beta}}$. Derive the sampling distribution of $\widehat{\beta}_k$.

ii. Let β_k be the kth element of $\boldsymbol{\beta}$. What is the sampling distribution of $(\widehat{\beta}_k - \beta_k)/se(\widehat{\beta}_k)$ where $se(\widehat{\beta}_k)$ is the square root of the estimated variance of the sampling distribution of $\widehat{\beta}_k$. You do not need to derive this result formally, but what facts are you using to arrive at your answer?

iii. Use your result from (ii) to construct a confidence interval for $\widehat{\beta}_k$ that has coverage $1 - \alpha$. Show formally that your interval has coverage $(1 - \alpha)$.