HW1 for STAT 818R Fall 2026, due Sunday 9/27 on ELMS

Eric Slud 9/17/2026

(1). We considered in class a least-squares problem in which $Y_i \sim \mathcal{N}(X_i^T \beta, \, \sigma^2)$ for $i = 1, \ldots, n$, for some scalar variance $\sigma^2 > 0$ and a full-rank $n \times p$ nonrandom design matrix X with rows X_i. The weighted least-squares estimator is defined as the $\mathbb{R}^p$ minimizer $\hat{\beta}$ of $(Y - X\beta)^T \, W \, (Y - X^T \beta)$ with respect to β, where W is a known $n \times n$ symmetric positive-definite matrix not depending on β and Y is the $\mathbb{R}^n$ vector with entries Y_i. This estimator $\hat{\beta}$ was shown in class to have variance (under the assumed homoscedastic normal-errors model)

$$\sigma^2 \, (X^T W X)^{-1} \, X^T W^2 X \, (X^T W X)^{-1} \tag{1}$$

If we want to compare the variance of the estimator with weight-matrix W versus the ordinary least-squares (OLS) estimator obtained by replacing W with the $n \times n$ identity matrix $I_{n \times n}$ then we are comparing the matrices

$$(X^T W X)^{-1} \, X^T W^2 X \, (X^T W X)^{-1} \qquad \text{versus} \qquad (X^T X)^{-1}$$

The assumed problem is to **prove directly by linear algebra** that

$$(X^T W X)^{-1} \, X^T W^2 X \, (X^T W X)^{-1} \, - \, (X^T X)^{-1}$$

is nonnegative definite and equal to 0 only when $W = c \cdot I_{n \times n}$ for a scalar constant c. If you want, you may assume that W is a diagonal matrix, but there is a straightforward linear-algebra proof that works for general symmetric positive-definite matrix W.

(2). Consider *iid* data $Z_i \sim \mathcal{N}(\mu, \, \frac{1}{2} \, \mu^2), \; i = 1, \ldots, n$.

(a). Find the large-sample variance of the MLE by finding the Fisher Information.

(b). Find the large-sample variance of the Method of Moments estimator based on the sample mean $\bar{Z}$ by regarding it as an Estimating Equation Estimator.

(c). Define a new estimator as suggested in class by minimizing a sum of squared discrepancies in the first and second moment equations,

$$\min_{\mu} \; \Big\{ (\bar{Z} \, - \, \mu)^2 \; + \; (n^{-1} \sum_{i=1}^{n} Z_i^2 \, - \, \frac{3}{2} \mu^2)^2 \; \Big\}$$

Find the large-sample variance of this estimator by regarding it as an Estimating Equation Estimator.

(d). Show that the variance in (b) is larger than the variance in (a), but that the large-sample variance in (c) is strictly between the (a) and (b) variances.

(3). Suppose that $Y_i = X_i^T \beta + v_i + e_i$ for $i = 1, \ldots, m$ as in the Fay-Herriot (FH)model, where $X_i \in \mathbb{R}^p$ are known constant vectors (with $\sum_{i=1}^n X_i^{\otimes 2}$ nonsingular) and $e_i \sim \mathcal{N}(0, D_i)$ with known positive constants D_i, and $v_i \sim \mathcal{N}(0, \sigma^2)$. Is there an optimal estimating-equation within the class

$$\sum_{i=1}^n \psi_i(X_i, Y_i, \beta, \sigma^2) = \sum_{i=1}^n W_i(\beta, \sigma^2) \begin{pmatrix} X_i(Y_i - X_i^T \beta) \\ (Y_i - X_i^T \beta)^2 - D_i - \sigma^2 \end{pmatrix}$$

where $W_i(\beta, \sigma^2)$ are positive definite $(p+1) \times (p+1)$ matrix-valued function twice continuously differentiable with respect to their arguments, not depending on Y_i ? (There would have to be some restrictions on the W_i's so that they satisfy some regularity properties when averaged across i; but ths question asks whether we should expect some choice to be optimal among a large class of such estimating equations, when the underlying FH model holds.)