Exercises November 13 **STA 4508S (Fall, 2018)**

1. Consider the simple linear regression model $Y_i = \beta_0 + \beta_1 x_i + \epsilon_i$, where ϵ_i are independent normal random variables with expected value zero and variance $\sigma_i^2 = \sigma^2(1 + \gamma x_i^2), i = 1, \ldots, n$. Simulate 1000 datasets of length $n = 50$ with parameters $\beta_0 = 1, \beta_1 = 1, \sigma^2 = 3, \gamma = 2$ and covariate x_i simulated from a $U(-1, 1)$.

   (a) Fit each dataset with a simple linear regression model (assuming $\gamma = 0$), and ~~compare~~compute the simulation mean and variance of $\hat{\beta}_1$ ~~to that computed from the fitted model with $\gamma = 0$~~.

   (b) Compare the true and estimated sandwich variance of $\hat{\beta}_1$ based on the Godambe information matrix to the naive estimate from (a) (from the regression output).

   (c) The true $\text{var}(\hat{\beta}_1)$ can be computed (tediously) from the appropriate element of G^{-1}, where $G(\cdot)$ is the Godambe information. Somewhat confusingly, this is a 3×3 matrix, since the fitted model has just 3 parameters, but it depends on $(\beta_0, \beta_1, \sigma^2, \gamma)$ (and these values are known since we are simulating). (Royal "we")

   The estimated value of the Godambe information is less clear, because we have no estimate of γ. However, if we compute the 3×1 score vector for each observation, say $U_i(\beta_0, \beta_1, \sigma^2)$ we can estimate $\text{E}(UU^{\text{T}})$ by

   $$\frac{1}{n}\sum_{i=1}^{n} U_i(\hat{\theta})U_i(\hat{\theta})^{\text{T}}. \tag{1}$$

   General least-squares theory shows that if $y \sim N(X\beta, \sigma^2 W)$, then $\hat{\beta}_{LS} = (X^{\text{T}}X)^{-1}X^{\text{T}}y$ has expected value β and variance

   $$\sigma^2(X^{\text{T}}X)^{-1}(X^{\text{T}}WX)(X^{\text{T}}X)^{-1},$$

   where W is a diagonal matrix whose entries can be estimated using $\hat{\epsilon}$. This agrees with what I got using (1).

   So together (a) and (b) ask you to compare the simulation variance of $\hat{\beta}_1$ with its true variance under the model, the estimated variance using (1), and the estimated variance from the regression fit $(\sigma^2(X^{\text{T}}X)^{-1})$.

2. Suppose we have an i.i.d. sample $y_1, \ldots, y_n$ from a model with density $f(y;\theta), \theta \in \mathbb{R}$, and we estimate θ by means of an estimating equation

$$g(y;\theta) = \sum_{i=1}^{n} g(y_i;\theta);$$

the estimator $\tilde{\theta}_g$ is defined as the solution to $g(y;\theta) = 0$, assuming the solution exists w.p.1. We assume that $g(y;\theta)$ is an unbiased estimating equation, i.e.

$$\mathrm{E}\{g(y;\theta)\} = \int g(y;\theta) f(y;\theta) dy = 0.$$

(a) Assuming further that g is differentiable with respect to θ, show that to a first order of approximation

$$\tilde{\theta}_g \doteq \theta - \frac{\sum_{i=1}^{n} g(y_i;\theta)}{\sum_{i=1}^{n} \partial g(y_i;\theta)/\partial\theta}.$$

(b) Apply the weak law of large numbers to the denominator, and the central limit theorem to the numerator, to conclude that

$$\sqrt{n}(\tilde{\theta}_g - \theta) \dot\sim N\{0, \sigma^2(\theta)\},$$

where

$$\sigma^2(\theta) = \frac{\mathrm{var}\{g(Y_1;\theta)\}}{\mathrm{E}\{-\partial g(Y_1;\theta)/\partial\theta\}^2}.$$