This is an independent take-home assignment. That is, you are *not* allowed to collaborate with each other, other students, faculty, or anyone else. You may use your notes, textbooks, or articles as an aid. You may *not* look at materials from previous versions of this course. Please show all your work as credit will not be given for undocumented answers. Questions of clarifications should be addressed directly to me or Tianchen Qian. State any critical assumptions. Please hand in the assignment by Friday, October 24 at 5pm. On your assignment, please sign the statement, "I have neither given nor received aid on this homework assignment." Good luck!

1. Let $X$ denote the concentration of a given drug and $Y$ is the continuous response. From previous experimentation with these types of drugs, it is known that

$$E[Y|X] = \theta_1 + \theta_2 \exp(-\theta_3 X) = \mu(X; \theta)$$

where $\theta = (\theta_1, \theta_2, \theta_3)$. It is also know that the variance of $Y$ given $X$ is proportional to the conditional mean. That is,

$$\text{Var}[Y|X] = \sigma^2 \mu(X; \theta)$$

Assume that we have $n$ i.i.d. copies of $(Y,X)$, i.e., $\{(Y_i, X_i) : i = 1, \ldots, n\}$.

a. Find the optimal generalized estimating equation for estimating $\theta$. Show that the unknown scale factor $\sigma^2$ does not affect the resulting estimates.

b. What is the asymptotic covariance matrix of $\hat{\theta}$. Assume that $\sigma^2$ is known and equal to $\sigma_0^2$.

Of course $\sigma^2$ is unknown and must be estimated from the data. A simple estimate for $\sigma^2$ is based on the application of method of moments. That is,

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} \frac{(Y_i - \mu(X_i; \hat{\theta}))^2}{\mu(X_i; \hat{\theta})}$$

c. Prove that $\hat{\sigma}^2$ is a consistent estimate of $\sigma^2$. State any regularity conditions that you find necessary to establish this result.

d. Suppose that the variance of $Y$ given $X$ was misspecified and $\text{Var}[Y|X] = V_{true}(X) \neq \sigma^2 \mu(X; \theta)$. What can we say about the estimates $\hat{\theta}$ derived from part a). What will be the asymptotic covariance matrix of $\hat{\theta}$. How would you estimate the asymptotic covariance of $\hat{\theta}$.

2. Let $X_1, \ldots, X_n$ be i.i.d., where $X_i = (W_i, Y_i, Z_i)$, $Z_i$ is an unobserved latent variable, $W_i$ is independent of $Y_i$ conditional on $Z_i$, $W_i$ given $Z_i$ is $\text{Exponential}(Z_i)$ and $Y_i$ given $Z_i$ is $\text{Exponential}(\gamma Z_i)$, and the distribution of $Z_i$, $f$, is unspecified. The goal is to use the observed data $\{(W_1,Y_1), \ldots, (W_n,Y_n)\}$ to draw inference about $\gamma$. Let

$$g(X,Y; \gamma) = \frac{W - \gamma Y}{W + \gamma Y}$$
a. Show that \( E_{\gamma,f}[g(W,Y;\gamma)] = 0 \) for all \( \gamma \) and \( f \).

b. Rigorously prove that the solution \( \hat{\gamma} \), the solution of

\[
\sum_{i=1}^{n} g(W_i,Y_i;\gamma) = 0
\]

is consistent and asymptotically normal. State any regularity conditions that you find necessary to establish your result.

3. Suppose that \((X_1, \ldots, X_n)\) are i.i.d., where the \( X_i \)'s are \( N(\theta, 1) \). Let

\[
g_c(x;\theta) = \begin{cases} 
-c & x - \theta \leq -c \\
 x - \theta & -c < x - \theta < c \\
c & x - \theta \geq c
\end{cases}
\]

Let \( \tilde{\theta}(c) \) denote the solution to \( \sum_{i=1}^{n} g_c(X_i;\theta) = 0 \). This is Huber’s estimator.

(a) Show that \( g_c(X;\theta) \) is an unbiased estimating function.

(b) Derive the asymptotic distribution of \( \tilde{\theta}(c) \)?

(c) Plot the asymptotic efficiency of \( \tilde{\theta}(c) \) relative to the MLE as a function of \( c \)?

4. Let \( X_1, \ldots, X_n \) be an i.i.d. sample of \( N(\theta, 1) \) random variables, where \( \theta \geq 0 \). Derive the asymptotic distribution of the MLE when \( \theta = 0 \). Does this result contradict any of the theorems presented in Chapter 4 of the course notes? Explain.

5. Let \( Z = (Z_1, Z_2)' \) be a bivariate normally distributed vector with mean zero and covariance matrix

\[
\begin{bmatrix}
\sigma_1^2 & \sigma_{12} \\
\sigma_{12} & \sigma_2^2
\end{bmatrix}
\]

Consider the Hilbert space of all one-dimensional measurable functions of \( Z \) with mean zero, finite variance, and covariance inner product. Let \( \mathcal{U} \) denote the linear subspace spanned by \( Z_2 \) and \( (Z_1^2 - \sigma_1^2) \), i.e. the space whose elements are

\[
\left\{ a_1(Z_1^2 - \sigma_1^2) + a_2Z_2 : a_1, a_2 \in \mathbb{R} \right\}
\]

(a) Find the projection of \( Z_1 \) onto the space \( \mathcal{U} \).

(b) Find the variance of the residual; i.e. \( var[Z_1 - \Pi[Z_1|\mathcal{U}]] \).