Statistics 582, Problem Set 3, revised

Wellner; 1/17/2018

Reading: van der Vaart, *Asymp. Stat.*, Section 5.3, pages 51-60;
Handout on Huber's $Z-$theorem, pages 1-10;
Chapter 5, sections 1-4; Ferguson, *Math Statist.*, chapter 1.

Due: Wednesday, January 24, 2018.

Reminder: Make up lecture 1, 9:30 - 10:20, Monday, 22 January, EEB 003

1. (vdV *Asymp. Stat.*, problem 14, page 83): Suppose that we observe a random sample from the distribution of (X, Y) in the following *errors in variables* model:

$$\begin{aligned} X &= Z + e \\ Y &= \alpha + \beta Z + f \end{aligned}$$

where (e, f) is bivariate normally distributed with mean 0 and covariance matrix $\sigma^2 I$ and is independent from the unobservable variable Z. In analogy to Example 5.26, construct a system of estimating equations for $\theta = (\alpha, \beta)$ based on a conditional likelihood, and study the properties of the corresponding estimators $\widehat{\theta}_n = (\widehat{\alpha}_n, \widehat{\beta}_n)$. In particular, what is the limiting distribution of $\sqrt{n}(\widehat{\theta}_n - \theta)$ under some reasonable assumptions about existence of moments of Z?

Hint: Condition on the unobserved $(Z_1, \ldots, Z_n) = (z_1, \ldots, z_n)$ and treat the z_i's as parameters in the model. Then the joint density $p(x_i, y_i) \equiv p(x_i, y_i; \alpha, \beta, z_i, \sigma^2)$ of (X_i, Y_i) is Gaussian. From there, proceed as follows:

(a) Show that the log-likelihood of the data $(X_1, Y_1), \ldots, (X_n, Y_n)$ is given by

$$l_n(\alpha, \beta, \underline{z}, \sigma^2) = -n \log(2\pi\sigma^2) - \frac{1}{\sigma^2} \left\{ \sum_{i=1}^n (Y_i - \alpha - \beta z_i)^2 + \sum_{i=1}^n (X_i - z_i)^2 \right\}.$$

(b) Maximize $l_n(\alpha, \beta, \underline{z}, \sigma^2)$ for fixed α, $\beta, \sigma^2)$ as a function of $\underline{z}$ to find

$$\begin{aligned} l_n^{prof,1}(\alpha, \beta, \sigma^2) &\equiv l_n(\alpha, \beta, \hat{\underline{z}}(\alpha, \beta), \sigma^2) \\ &= -n \log(2\pi\sigma^2) - \frac{1}{\sigma^2(1+\beta^2)} \sum_{i=1}^n (Y_i - \alpha - \beta X_i)^2. \end{aligned}$$

(c) Maximize $l_n^{prof,1}(\alpha, \beta\hat{\underline{z}}(\alpha, \beta), \sigma^2)$ for fixed β, σ^2 to find

$$\begin{aligned} l_n^{prof,2}(\beta, \sigma^2) &= l_n^{prof,1}(\hat{\alpha}(\beta), \beta, \sigma^2) \\ &= -n \log(2\pi\sigma^2) - \frac{1}{\sigma^2(1+\beta^2)} \sum_{i=1}^n (Y_i - \overline{Y} - \beta(X_i - \overline{X}))^2. \end{aligned}$$

(d) Maximize $l_n^{prof,2}$ with respect to β to find $\hat{\beta}$. You should find that:

$$\begin{aligned} \hat{z}_i &\equiv \hat{z}_i(\alpha, \beta) = X_i + \frac{(Y_i - \alpha - \beta X_i)\beta}{1+\beta^2}, \quad i = 1, \ldots, n, \\ \hat{\alpha} &\equiv \hat{\alpha}(\beta) = \overline{Y} - \beta \overline{X}, \\ \hat{\beta} &= \frac{S_{YY} - S_{XX} + \sqrt{(S_{YY} - S_{XX})^2 + 4S_{XY}^2}}{2S_{XY}} \end{aligned}$$

where

$$S_{XX} = n^{-1}\sum_{i=1}^{n}(X_i - \overline{X})^2, \qquad S_{YY} = n^{-1}\sum_{i=1}^{n}(Y_i - \overline{Y})^2,$$
$$S_{XY} = n^{-1}\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y}).$$

(e) Now study the asymptotic properties of $(\hat{\alpha}_n, \hat{\beta}_n)$.
This is sometimes called *orthogonal regression*, since the fitted line minimizes orthogonal distances from observed points to the regression line. If instead of equal variances for the errors we had assumed $\sigma_\epsilon^2/\sigma_f^2 = \delta$, then the solution becomes

$$\begin{aligned}
\hat{z}_i &\equiv \hat{z}_i(\alpha, \beta) = X_i + \frac{(Y_i - \alpha - \beta X_i)\beta}{\delta + \beta^2}, \quad i = 1, \ldots, n, \\
\hat{\alpha} &\equiv \hat{\alpha}(\beta) = \overline{Y} - \beta\overline{X}, \\
\hat{\beta} &= \frac{S_{YY} - \delta S_{XX} + \sqrt{(S_{YY} - \delta S_{XX})^2 + 4\delta S_{XY}^2}}{2S_{XY}}.
\end{aligned}$$

This is called *Deming regression*. For semiparametric (*structural*) versions of these models, see Bickel and Ritov (1987) and van der Vaart (1996).

2. (a) (vdV *Asymp. Stat.*, problem 15, page 83): In Example 5.27, for what point is the least squares estimator consistent if we drop the the condition that $E(e|X) = 0$? Derive an (implicit) solution in terms of the function $E(e|X)$. Is it necessarily θ_0 if $E(e) = 0$?
(b) Investigate the limit distribution of $\sqrt{n}(\widehat{\theta}_n - \theta_0)$ under the assumption $E(e|X) = 0$.

3. Consider the zero-inflated Poisson distribution p_θ as described in Example 3 of the handout on M- and Z- theorems. Suppose that $X_1, \ldots, X_n$ i.i.d. p_θ are observed.
(a) Set up alternative estimating equations for $\theta = (\gamma, \lambda)$ where $\gamma \in [0, 1]$ and $\lambda > 0$ based on $g_1(x) = x$ and $g_2(x) = x^2$. Express your alternative estimator $\hat{\theta}_n = (\hat{\gamma}_n, \hat{\lambda}_n)$ of θ explicitly in terms of the first and second moments, $\overline{X}_n$ and $\overline{X^2}_n$, of the data, and show that your estimators are consistent when the model holds.
(b) Use Huber's Z-theorem to show that $\sqrt{n}(\hat{\theta}_n - \theta_0) \to_d N_2(0, \Sigma)$ and give the form of Σ.
(c) What happens if the X_i's are i.i.d. $p \notin \mathcal{P} = \{p_\theta : \theta \in \Theta\}$? Describe the parameter $\theta(P)$ to which $\hat{\theta}_n$ converges in probability and use Huber's theorem to establish a limit theorem for $\sqrt{n}(\hat{\theta}_n - \theta(P))$ in this case.

4. Let X be a random variable with distribution function F having finite first moment: $E|X| < \infty$.
(a) Show that $f(b) \equiv E|X - b|$ is minimized by $b =$ any median of the distribution F of X. [A median m of F is any value satisfying $F(m) = P(X \leq m) \geq 1/2$ and $1 - F(m-) = P(X \geq m) \geq 1/2$; see Lehmann and Casella, TPE, page 62, problems 1.7 and 1.8.]
(b) For $0 < \tau < 1$, let $\rho_\tau(x) = x(\tau - 1_{(-\infty,0)}(x))$. Consider minimizing

$$M_\tau(\theta) = E\rho_\tau(X - \theta)$$

with respect to θ. Show that the solution $\theta_0 = \theta_0(F)$ is given by the $\tau-$th quantile of F: $\theta_0(F) = F^{-1}(\tau)$.