

Statistics 582, Problem Set 3

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 $\hat{\theta}_n = (\hat{\alpha}_n, \hat{\beta}_n)$. In particular, what is the limiting distribution of $\sqrt{n}(\hat{\theta}_n - \theta)$ under some reasonable assumptions about existence of moments of Z ?

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 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 $\sqrt{n}(\hat{\theta}_n - \theta_0)$ 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, \dots, 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, \bar{X}_n and \bar{X}_n^2 , of the data, and show that your estimators are consistent when the model holds.
 - (b) Use Huber's Z-theorem to show that $\sqrt{n}(\hat{\theta}_n - \theta_0) \rightarrow_d N_2(0, \Sigma)$ 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. What other methods do you have available 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 τ -th quantile of F : $\theta_0(F) = F^{-1}(\tau)$.