Problem Set 3:

Bootstrap, Quantile Regression and MCMC Methods

MIT 14.385, Fall 2007

Due: Wednesday, 07 November 2007, 5:00 PM

1 Applied Problems

Instructions:

The page indications given below give you an upper bound on what you should write. Answers should be very brief. Hand in clear, annotated code. You are strongly encouraged to work in groups, but the write-ups should be individual.

1.1 Bootstrap

1. Explain briefly what the bootstrap is, why it works when it works and how it can fail. (1/2 page max, informal reasoning is fine)

2. Find an empirical paper that uses the bootstrap. Explain why they use the bootstrap. (1/4 page max) Write "pseudo-code" (a brief outline of an algorithm) for how you would reproduce the bootstrap used in this paper. (1/2 page max)

3. What is the bootstrap bias correction method? How does it work? State the pseudo-code for bias correction. (1/2 page max, informal reasoning is fine; hint: read Horowitz)

4. Two-step estimators have messy standard errors, which makes bootstrap inference appealing. Consider the censored regression 2-step estimator of PSet 2, Q1.b.iii. Bootstrap the standard errors. Compare these standard errors to those generated for you by the Stata canned command.

1.2 Quantile Regression

1. Theory. Briefly answer the following questions.

   Fundamentals:

   (a) (1/4 page max) Consider the location-scale model

   $$Y = X'\alpha + (X'\gamma)\, u$$

   where u is independent of x and has distribution function $F_u(u)$ Write down the conditional quantile function of y given x in the form

   $$Q_Y(\tau|x) = x'\beta(\tau)$$

   Characterize $\beta(\tau)$ as a function of τ (e.g. monotonicity if any ...). How does the answer change if $(x'\gamma) = 1$?

   (b) (1/4 page max) Show that $E[Y|x] = \int Q_Y(\tau|x)\, d\tau$ is of the form $x'\beta$, where β is what?

   (c) (1/4 page max) Consider a general linear quantile model, $Q_Y(\tau|x) = x'\beta(\tau)$ corresponds to the random coefficient model $Y = X'\beta(U)$. Can we have slope coefficients $\beta(\tau)$ that behave non-monotonically in τ? Recall Doksum's quantile treatment effect example.

   (d) (1/2 page max) Using (a)-(b), give a brief conceptual evaluation of the linear quantile regression model. Compare it with the classical linear model, and state the advantages and disadvantages of the two. Refer to at least one empirical example discussed in class where the difference in models is important. E.g. refer to Doksum's guinea pig example.

   (e) (1 page max) Quantile equivariance to monotone transformations

   Suppose the conditional quantile function of y given x is in the form

   $$Q_Y(\tau|x) = x'\beta(\tau)$$

   Now consider the following models of data transformation

i. $y^* = f(y) = \max\{y, 0\}$

ii. $y^* = f(y) = 1\{y > 0\}$

iii. $y^* = f(y) = \exp(y)$

For each of the above, write down the conditional quantile function for the trasnformed variable conditional on x:

$$Q_{Y^*}(\tau|x) = f(x'\beta(\tau))$$

Hint: quantile models have a remarkable property — equivariance to monotone transformations, which we (should) observe in these examples. (Useful reference: `http://www.econ.uiuc.edu/~roger/research/rq/rq.pdf` + lecture notes).

Is the same property true of the conditional mean function, i.e. is $E[Y^*|x] = f[E[Y|x]]$? Why is this a problem?

(f) (1/2 page) For each of the above three cases write down an estimators of $\beta(\tau)$ that use data on (y^*, x) only. (Hint: Case (i) generates the censored quantile regression estimator, case (ii) the maximum score / perception / single-layer neural network estimator, and case (iii) a nonlinear quantile regression estimator.)

(g) (1/4 page maximum) Summarize the findings above very briefly, and say why they are important for applications to duration models and censored regression models.

2. (1 page, excluding figures.). Empirical: Reproduce or closely replicate Figures 1, Table 1, then Figure 2 of Koenker's "Vignette" (posted in Recitation Materials section on MIT Server alternatively Roger Koenker's website). You may use any software you like, but is probably just easier to use R (which is a great statistical freeware). As usual, you should briefly explain what the canned commands are doing. The Engel data are available from within R or on the MIT Server.

3. For this question, use the data in "Penn46.ascii". See Section 5 of Koenker and Billias, "Quantile Regression for Duration Data," Empirical Economics, 2001, for a description of the

program and the data. For part (a), review W. Newey's lecture note on duration models or Wooldridge 20.1-20.2.

(a) (1 page) Conventional duration models usually boil down to an empirical relation of the following kind:

$$h(T_i) = x_i'\beta + u_i \tag{1}$$

where T_i is a "survival time" (here, the duration of an unemployment spell), $h()$ is monotonic and u_i is iid with CDF F_u.

(i) Show that if we assume a *Cox proportional hazard model,* we can write the duration of unemployment in the form in (a) above with

$$h(T) = \log \Lambda_0(T)$$

where $\Lambda_0(t)$ is the integrated baseline hazard.

(ii) Show that in the special case of a *Weibull baseline hazard*, we have

$$\log \Lambda_0(T) = \gamma \log T - \alpha$$

(iii) Combine your results from (i) and (ii) to show that, under the assumptions of Cox proportional hazard *and* Weibull baseline hazard, the log of the unemployment duration satisfies the following equation:

$$\log T_i = x_i'\beta + u_i$$

This is a special case of the model called the *Accelerated Failure Time (AFT) model.* (Note: AFT models do not make parametric assumptions on the error term).

(b) (1/2 page) Show that an implication of the model (1) is that the covariates only shift the *location* of the distribution, not the scale, i.e. the conditional quantile function of

$h(T)$ is

$$Q_{h(T)}(\tau|x) = x'\beta + F_u^{-1}(\tau)$$

for $\tau \in (0,1)$.

Show that we can write the conditional quantile function of T as $Q_T(\tau|x) = h^{-1}\left(x'\beta + F_u^{-1}(\tau)\right)$ Do covariates affect the location of T only (but not the scale etc.?)

(c) (1/2 page) Now consider a quantile regression alternative to (1). Explain how quantile regression can be used to estimate more general models of the form $Q_{h(T)}(\tau|x) = x'\beta(\tau)$, and how if we wanted, we could recover an estimate of

$$Q_T(\tau|x) = h^{-1}(x'\beta(\tau))$$

What are pros and cons of quantile regression approach relative to the approach outlined in (a) and (b)?

(d) (1 page, excluding graphs) Using the Penn data, estimate the AFT model by maximum likelihood (using the Weibull distribution — state what the likelihood function is) and by OLS, using the following regressors:

indicators for the 5 treatment groups, with treatments 4 and 6 pooled; indicators for female, black, and hispanic respondents; number of dependents, with 2 indicating two or more dependents; indicators for the 5 quarters of entry to the experiment; indicator for whether the claimant "expected to be recalled"; indicator of whether the respondent was "young" — less than 35 — or "old" — greater than 54; indicator for whether claimant was employed in the durable goods industry; indicator for whether the claimant was registered in a low employment district (Coatesville, Reading or Lancaster). (All these are coded for you in the data set and it should be clear from the variable names what to use.)

Present your results in a neat, clear format that is easy to read. Give a brief discussion of the key results.

(e) (1 page, excluding graphs) Again using the Penn data and the same regressors, estimate the model $Q_{\log(T)}(\tau|x) = x'\beta(\tau)$ by quantile regression for $\tau = \{0.20, 0.25, \ldots, 0.75, 0.80\}$. Present your results graphically.

(f) (1/4 page) What do you make of results in (d) and (e)?

1.3 Bayesian and Quasi-Bayesian Estimators

1. Explain briefly what these estimators do. Give a simple one page explanation.

2. Explain what is (quasi) posterior mean, median, quantile. Are posterior mean and medians inferior estimators compared to the extremum estimates? Are they consistent, asymptotically normal? Do posterior quantiles provide valid inference when generalized information equality holds (satisfied for example when GMM function has optimal weighting matrix). (1/2 page max)

1.4 Introduction to MCMC:

1. Consider doing IV median regression using GMM criterion function for the model:

   $y = \beta x + u, \ \ u \sim N(0, 1),$

   where $x = z + v, \ \ v \sim N(0, 1)$ and $z \sim N(0, 1)$. Here z is the instrument, x is the endogenous variable, and y is outcome. Suppose the true value of β is 0. Suppose the sample size is 100.

   Write an MCMC (Metroppolis) algorithm that produces a posterior mean and median estimates of β. Compute also the confidence intervals for β based on posterior quantiles. Report your estimates and confidence intervals. Briefly explain (one paragraph) how you would set up the objective function to use in the quasi-bayes estimation. Briefly explain (a paragraph) what the code is doing. Below I provide the sample code for this problem, which is written in R. Feel free to use it, but make sure to explain it. Print out the graphs as well.

   Also, I should note that for future work, that there are several professional implementations of the Metropolis algorithm, for example, you can find one by Charles Geyer at his homepage at the stats department of the University of Minnesota (you do not need it for this problem).

2 Theory Problems

If you do this option you can skip the bootstrap question (1.1) and the empirical portion of the quantile regression question (but still do the non-empirical parts). You also need to do Bayesian and the MCMC questions.

1. Read parts of Andrew's Handbook chapter (p. 2249–58) that deals with stochastic equicontinuity and its use for deriving the asymptotic properties of estimators. Present the Andrews argument for the asymptotic normality. Explain the role of stochastic equicontinuity in this derivation. Explain how one can verify the stochastic equicontinuity via bracketing (see p. 2276–).[1]

2. Use Andrew's approach to characterize the asymptotic distribution of quantile regression estimator. Rigorously verify the stochastic equicontinuity of a relevant empirical process, using the bracketing entropy approach (you may also use symmetrization approach if you like).

[1] Andrews, Donald W. K. "Empirical Process Methods in Econometrics." In *Handbook of Econometrics*, vol. 4, edited by Daniel McFadden and Robert Engle. Amsterdam, The Netherlands: North-Holland Publishing Co., 1994, chap. 37, pp. 2247-2294. ISBN: 9780444887665.