University of Illinois	**Econ 574**	Department of Economics
Spring 2008	**Final Exam**	Professor Roger Koenker

This is a take home final exam. It will be available on the web at 9am Friday May 2 and it is due on Friday, May 9 by 5pm. You can put it in my mailbox in 484 Wohlers Hall. The two questions are meant to be representive of the kind of questions that you might face later in research, and for which (I hope) the course provided some general preparation.

1. (Slicing Dicing and Stein's Lemma.) This question was inspired by the talk by Ker-Chau Li at the 2008 Bohrer Workshop.

   (a) Suppose that $u \sim U[0,1]$ and G is a continuously differentiable function with derivative, g, on $[0,1]$. Show that

   $$Eg(u) = G(1) - G(0)$$

   (b) Now suppose that you would like to compute

   $$M = \int_0^1 (\log(1/x))^{1/2} dx$$

   use the result in (a.) to approximate M by the sample mean

   $$M_n = n^{-1} \sum_{i=1}^{n} \sqrt{\log(1/u_i)}$$

   and evaluate the quality of the approximation for several sample sizes. Do the results conform to what you are led to expect by the CLT?

   (c) Suppose instead that $Z \sim \mathcal{N}(0,1)$ and G is continuously differentiable with derivative g on the whole real line. Show (Stein's Lemma) that:

   $$Eg(Z) = \text{Cov}(Z, G(Z)).$$

   And illustrate numerically with the simple example $g(z) = \cos(z)$.

   (d) A slight generalization of the foregoing result, usually attributed to Brillinger, is

   Suppose (X, Y) are jointly normal with G, g as in Stein's lemma,

   $$\text{cov}(G(X), Y) = \frac{\text{cov}(X, Y)}{V(X)} Eg(X)$$

   Prove. Hint: Use the fact that,

   $$\text{Cov}(G(X), Y) = \text{cov}(G(X), E(Y|X))$$

   and Stein's Lemma.

(e) Now, (finally!) we are getting close to modern developments in econometrics, consider the transformation model

$$E(Y|X = x) = h(x^\top \beta) \equiv m(x)$$

where h is a monotone (increasing) transformation. We would like to be able to estimate β without prior knowledge of h. That is, we are interested in the gradient vector

$$\begin{aligned} \nabla m(x) &= \left(\frac{\partial E(Y|X = x)}{\partial x_i} \right) \\ &= h'(x^\top \beta) - \beta \end{aligned}$$

Obviously, $\nabla m(x)$ depends upon where we evaluate the function

$$\gamma(X) = h'(x^\top \beta)$$

but suppose that we are satisfied to estimate the *average derivative*

$$E_X \nabla m(x) = \int \nabla m(x) f_X(x) dx$$

where f_X denote the joint density of X. Show that if X happens to be jointly normally distributed then this average derivative is consistently estimated *up to scale* by the least-squares regression coefficients of Y on X. That is, for some scalar κ,

$$E_X \nabla m(X) = \kappa \Sigma_{XX}^{-1} \Sigma_{XY}$$

where

$$\mathrm{Cov}(X, Y) = \begin{pmatrix} \Sigma_{XX} & \Sigma_{XY} \\ \Sigma_{XY}^\top & \Sigma_{YY} \end{pmatrix}.$$

Epilogue: There is an extensive literature dealing with the application of these results to the diagnosis of non-linearities in regression modeling. Unfortunately the normality of X is quite implausible in most applications. but similar results can be extended to elliptically symmetric X But I will resist the temptation to expand this line of questions further.

2. (Non-parametric Regression) This question was motivated by an R-help inquiry I had a few months ago from a geneticist, and uses one of the data sets that he provided. Consider the following pair of scatter plots:

The panel on the left illustrates a complete $x - y$ scatter plot of 311 points; the panel on the right illustrates a plot of the subset of points with $|y_i| < 1$. The objective is to find a simple and fairly automatic method of fitting a piecewise linear "curve" to the points that is not too badly influenced by the outliers.

Data is available in the file `Data.d` and the R-code for the Figure above is also available as `fig1.R`.

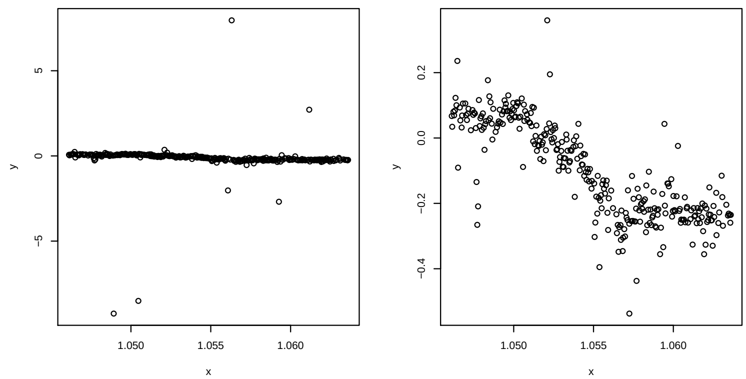

Figure 1: Scatterplots with and without Outliers

(a) The function `smooth.spline` in base R computes the classical $\mathcal{L}_2$ smoothing spline of Reinsch and Wahba. Try fitting the full and reduced data with this function using the default selection of λ in both cases, and then experiment with alternative λ-selection strategies.

(b) Repeat the exercise in (a.), but instead of using `smooth.spline` use the quantreg function

```
f  <- rqss(y ~ qss(x, lambda=.5))
```

Note that you can overplot the fitted function on the scatterplot with the command, e.g.

```
plot(f, col="red").
```

Experiment with the λ choice and also compare using the full and reduced samples. Compare and contrast *briefly* with the results in (a.).

(c) Explore more formally automatic procedures for choosing λ based on the AIC numbers returned by the `rqss` function.