University of Illinois
Spring 2002

Econ 476
Final Exam Answers

Department of Economics
Professor Roger Koenker

1. Part (a.) of this question seemed to be more difficult than I expected it to be, perhaps I didn't emphasize sufficiently in lecture the difficulties associated with the Wald approach.

   (a) Write,

$$\hat{\beta}_W = \frac{\bar{y}_R - \bar{y}_L}{\bar{z}_R - \bar{z}_L} = \beta - \beta \frac{\bar{v}_R - \bar{v}_L}{\bar{z}_R - \bar{z}_L} + \frac{\bar{u}_R - \bar{u}_L}{\bar{z}_R - \bar{z}_L}$$

The last term is $o_p(1)$ due to the $\perp\!\!\!\perp$ of u's and z's, but the second terms creates problems. Consider*

$$\begin{aligned} E\bar{v}_R &= E(v|z - \mu > 0) \\ &= \frac{\sigma_{vz}}{\sigma_z^2} E(z|z - \mu > 0) \\ &= \frac{\sigma_v^2}{\sigma_z^2} E(z|z - \mu > 0) \end{aligned}$$

Note that we rely crucially on the multivariate normality of (v, z) to write the conditional expectation

$$E(v|z) = \frac{\sigma_{vz}}{\sigma_z^2} z$$

and clearly, $\sigma_{vz} = E(vz) = E(v(x + v)) = \sigma_v^2$. Similarly,

$$E v_L = \frac{\sigma_v^2}{\sigma_z^2} E(z|z - \mu < 0)$$

Furthermore,

$$\begin{aligned} E\bar{z}_R &= E(z|z - \mu > 0) \\ E\bar{z}_L &= E(z|z - \mu < 0). \end{aligned}$$

So

$$\begin{aligned} \hat{\beta}_W &= \beta \left(1 - \frac{\sigma_v^2}{(\sigma_x^2 + \sigma_v^2)} \right) + o_p(1) \\ &= \beta \left(\frac{\sigma_x^2}{\sigma_x^2 + \sigma_v^2} \right) \end{aligned}$$

This bias factor is the same as for the LS estimator. When μ is estimated, we have $\hat{\mu} \to \mu$ and the conditional expectations tend to the same limits given above, so the asymptotic bias is the same as in the fixed μ case.

*We (provisionally) treat $\hat{\mu}$ as fixed.

(b) When $\sigma_v^2 = 0$ we have,

$$\hat{\beta}_W - \beta = \frac{\bar{u}_R - \bar{u}_L}{\bar{z}_R - \bar{z}_L}$$

and

$$\begin{aligned} E(\hat{\beta}_W - \beta)^2 &\to \frac{E(\bar{u}_R - \bar{u}_L)^2}{E(\bar{z}_R - \bar{z}_L)^2} = \frac{\sigma_u^2/n}{E(z_R^2 - 2z_R z_L + z_L^2)} \\ &= \frac{\sigma_u^2/n}{4E\bar{z}_R^2} = \frac{\sigma_u^2/n}{4(4\sigma_x^2/2\pi)} \\ &= \frac{\sigma_u^2}{\sigma_x^2}\frac{\pi}{8} \end{aligned}$$

thus ARE is

$$\text{ARE } (\hat{\beta}_W, \hat{\beta}_{OLS}) = \frac{\text{Avar } (\hat{\beta}_W)}{\text{Avar } (\hat{\beta}_{OLS})} = \frac{\pi}{8} \cong .40$$

(c) Obviously, in the normal case the Wald estimator isn't very successful as a bias reduction device. In non-normal cases there is somewhat more hope that it could be successful, but given the cost in variability as well there is not much to suggest, based on the normal theory, that it would be advantageous.

2. This question seemed to cause less trouble. I've included my R code and the resulting table.

```
#Spring 2002 Econ 476 Final Exam Question 2
lik <- function(theta,x){
#log likelihood function for iid cauchy location model scale ==1
sum(-log(1/(1+(x-theta)^2)))
}
onestep <- function(theta,x){
#onestep mle for the iid cauchy location model scale == 1
grad <- sum(2*(x-theta)/(1+(x-theta)^2))
hess <- sum(+4*(x-theta)^2/(1+(x-theta)^2)^2 - 2/(1+(x-theta)^2))
theta - grad/hess
}
R <- 1000
ns <- c(20,40,100)
A <- array(0,c(4,3,R))
for(j in 1:3){
        n <- ns[j]
        for(i in 1:R){
                x <- rt(n,1)
                A[1,j,i] <- mean(x)
                A[2,j,i] <- median(x)
```

```
                A[3,j,i] <- onestep(median(x),x)
                A[4,j,i] <- nlm(lik,median(x),x=x)$estimate
                }
        }
a <- apply(A^2,c(1,2),mean)

#make latex table for exam.R output
dimnames(a) <- list(c("mean","median","onestep","mle"),c("n=20","n=40","n=100"))
caption <- "Mean Squared Errors for 4 estimators of location for iid Cauchy
        observations and 3 sample sizes:  1000 replications"
tab <- format(round(a,3))
latex.table(tab ,caption=caption)
```

This yields the following results.

tab	n=20	n=40	n=100
mean	7850.153	2653.092	381.843
median	0.148	0.069	0.025
onestep	0.137	0.059	0.020
mle	0.129	0.059	0.020

Table 1: Mean Squared Errors for 4 estimators of location for iid Cauchy observations and 3 sample sizes: 1000 replications

3. The log likelihood is

$$\begin{aligned} l(\mu) &= -\frac{n}{2}\log(2\pi) - n\log\mu - \frac{1}{2}\sum\left(\frac{y_i-\mu}{\mu}\right)^2 \\ \nabla l &= -\frac{n}{\mu} + \sum\left(\frac{y_i-\mu}{\mu}\right)\left(\frac{1}{\mu}+\frac{y_i-\mu}{\mu^2}\right) = 0 \\ &\Rightarrow n^{-1}\sum(y_i-\mu)y_i = \mu^2 \\ &\Rightarrow \mu^2 + \bar{y}\mu - S_n = 0 \end{aligned}$$

where $\bar{y} = n^{-1}\sum y_i$ and $S_n = n^{-1}\sum y_i^2$.

$$\hat{\mu} = \frac{-\bar{y} \pm \sqrt{\bar{y}^2 + 4S_n}}{2}$$

Note by KSSN, $\bar{y} \to \mu$ and $S_n \to \mu^2 + \mu^2 = 2\mu^2$. Thus,

$$\hat{\mu} \to \frac{-\mu \pm \sqrt{9\mu^2}}{2} = \{\mu, -2\mu\}$$

Obviously, we want to choose the positive root. To find variance of mle, write

$$\begin{aligned}
\nabla l &= -\frac{n}{\mu} + \frac{1}{\mu^3}\sum y_i(y_i - \mu) \\
\nabla^2 l &= \frac{n}{\mu^2} - \frac{3}{\mu^4}\sum y_i(y_i - \mu) - \frac{1}{\mu^3}\sum y_i \\
E\nabla^2 l &= \frac{n}{\mu^2} - \frac{3}{\mu^4}[2n\mu^2 - n\mu^2] - \frac{n\mu}{\mu^3} \\
&= n\left[\frac{1}{\mu^2} - \frac{3}{\mu^2} - \frac{1}{\mu^2}\right] = -\frac{3n}{\mu^2}
\end{aligned}$$

so the mle satisfies $\sqrt{n}(\hat{\mu}-\mu) \rightsquigarrow \mathcal{N}(0, \mu^2/3)$ whereas $\sqrt{n}(\bar{y}-\mu) \rightsquigarrow \mathcal{N}(0, \mu^2)$ which requires 3 times as many observations to achieve the same precision.