UC Berkeley **Fall Semester 2004**

Economics 241A **Guido Imbens**

PROBLEM SET I, Part 4

SOLUTIONS

1. (a) Suppose that $X_1, \ldots, X_N$ are independent and identically distributed random variables with normal distributions with mean μ and variance σ^2. Describe an information matrix test using the $(1,2)$ element of the information matrix. What alternatives is this test powerful against?

The log of the density is

$$\ln f(x; \mu, \sigma^2) = -\frac{1}{2}\ln(2\pi) - \frac{1}{2}\ln\sigma^2 - \frac{1}{2\sigma^2}(x-\mu)^2.$$

The derivative with respect to μ is

$$\mathcal{S}_\mu(x, \mu, \sigma^2) = \frac{1}{\sigma^2}(x-\mu).$$

The derivative with respect to σ^2 is

$$\mathcal{S}_{\sigma^2}(x, \mu, \sigma^2) = -\frac{1}{2\sigma^2} + \frac{1}{2\sigma^4}(x-\mu)^2.$$

The second (cross) derivative is

$$\mathcal{H}_{12}(x, \mu, \sigma^2) = -\frac{1}{\sigma^4}(x-\mu).$$

Hence the information matrix test corresponds to first constructing the matrix W with ith row equal to

$$W_{i\cdot} = \Big(\mathcal{S}_\mu(x_i, \mu, \sigma^2) \;\; \mathcal{S}_{\sigma^2}(x_i, \mu, \sigma^2) \;\; \mathcal{H}_{12}(x_i, \mu, \sigma^2) + \mathcal{S}_\mu(x_i, \mu, \sigma^2)\cdot\mathcal{S}_{\sigma^2}(x_i, \mu, \sigma^2)\Big)$$

$$= \left(\frac{1}{\sigma^2}(x-\mu) \quad -\frac{1}{2\sigma^2} + \frac{1}{2\sigma^4}(x-\mu)^2 \right.$$
$$\left. -\frac{1}{\sigma^4}(x-\mu) + \frac{1}{\sigma^2}(x-\mu)\cdot\left(-\frac{1}{2\sigma^2} + \frac{1}{2\sigma^4}(x-\mu)^2\right)\right).$$

Then the IM test-statistic is equal to

$$\iota' W (W'W)^{-1} W' \iota,$$

which under the null has a Chi-squared distribution with one degree of freedom.

Note that the third row in W has elements depending on the third moment of $(X-\mu)$. Hence the test focuses on the skewness of the distribution of X, which of course should be zero for the normal distribution.

Also note that one cannot do the information matrix test based on the first element. This element would be equal to

$$\mathcal{H}_{11} + \mathcal{S}_\mu^2 = -\frac{1}{\sigma^2} + \frac{1}{\sigma^4}(x-\mu)^2,$$

which is equal to twice the score for σ^2.

(b) Suppose that $X_1, \ldots, X_N$ are independent and identically distributed random variables with Poisson distributions with parameter λ. Give the form of the information matrix test. What feature of the Poisson distribution is being tested?

The probability function for the Poisson distribution with parameter λ is equal to

$$f(x;\lambda) = \frac{\lambda^x \exp(-\lambda)}{x!}.$$

The log of the probability function is

$$\ln f(x;\lambda) = x \ln \lambda - \lambda - \ln(x!).$$

The score is

$$\mathcal{S}(x;\lambda) = \frac{x}{\lambda} - 1.$$

The Hessian is

$$\mathcal{H}(x;\lambda) = -\frac{x}{\lambda^2}.$$

Hence the W matrix of the IM test has ith row equal to

$$\begin{aligned} W_{i\cdot} &= \begin{pmatrix} \mathcal{S}(x_i;\lambda) & \mathcal{H}(x_i;\lambda) + \mathcal{S}(x_i;\lambda)^2 \end{pmatrix} \\ &= \begin{pmatrix} \frac{x}{\lambda} - 1 & \frac{x^2 - 2x\lambda + \lambda^2 - x}{\lambda^2} \end{pmatrix}. \end{aligned}$$

Essentially the test compares the variance of X to its mean. Under the Poisson model the two are equal.

2. Use the matlab data set NLS2.MAT which is posted on the website. It contains data on 935 observations on five variables: a binary indicator for finishing high school, two

tests scores, iq and kww, and mother's and father's education. We will look at logistic regression models with as the outcome variable the highschool indicator. The model we consider is

$$\Pr(\text{highschool} = 1|\text{kww}, \text{iq}, \text{med}, \text{fed})$$

$$= \frac{\exp(\beta_0 + \beta_1 \cdot \text{kww} + \beta_2 \cdot \text{iq} + \beta_3 \cdot \text{med} + \beta_4 \cdot \text{fed})}{1 + \exp(\beta_0 + \beta_1 \cdot \text{kww} + \beta_2 \cdot \text{iq} + \beta_3 \cdot \text{med} + \beta_4 \cdot \text{fed})}.$$

(a) Show that the log likelihood function at $\beta_0 = 2$, $\beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$ is equal to -294.6777.

(b) Calculate the first derivatives of the log likelihood function analytically at these values for the parameters. Compare them to numerical derivatives using $c = 0.001$.

See Table 1

Table 1: ANALYTICAL AND NUMERICAL FIRST DERIVATIVES AT $(2, 0, 0, 0, 0)$ USING $c = 0.001$

variable	analytic f.d.	numerical f.d
intercept	23.4547	23.4057
kww	39.5692	39.5316
iq	136.5924	136.4805
med	490.3028	484.8204
fed	399.4853	395.0150

(c) Estimate the parameters by maximum likelihood, using the DFP algorithm. Stop when the sum of the absolute values of the first derivatives is less than 0.001. Report the usual statistics for each iteration.

See matlab output

(d) Estimate the standard errors using the outer product of the first derivatives. Compare them to standard errors based on the A_k matrix, using the fact that the A_k matrix approximate the second derivative of the objective function.

See Table 2

Table 2: ML ESTIMATES

	unrestricted ml estimates				restricted ml estimates			
variable	est	se (fd)	se (A_k)	se (sd)	est	se (fd)	se (A_k)	se (sd)
intercept	1.9045	(0.3136)	(0.2955)	(0.3069)	1.5730	(0.2862)	(0.2595)	(0.2758)
kww	0.5266	(0.1796)	(0.1700)	(0.1712)	0	–	–	–
iq	0.5584	(0.1033)	(0.0906)	(0.0926)	0.6661	(0.0935)	(0.0860)	(0.0864)
med	0.1078	(0.0311)	(0.0311)	(0.0325)	0.1118	(0.0313)	(0.0309)	(0.0318)
fed	0.0131	(0.0331)	(0.0303)	(0.0303)	0.0127	(0.0331)	(0.0294)	(0.0298)

(e) Now estimate the parameters given $\beta_1 = 0$ by maximum likelihood, using the DFP algorithm. No need to report any statistics for each iteration, just the total number of iterations.

The number of iterations was 9. The output is in Table 2.

(f) Estimate the standard errors using the outer product of the first derivatives.

(g) Next we consider a Hausman-Wu test of the hypothesis $\beta_1 = 0$ by comparing the restricted and unrestricted estimates of β_3. What is the difference in the restricted and unrestricted estimates?

The difference in estimates is $\hat{\beta}_{3r} - \hat{\beta}_{3u} = -0.0040$.

(h) What are the two variances for $\hat{\beta}_3$ from (d) and (f)?

The variance of the unrestricted estimate, based on evaluating the variance at the unrestricted estimates is $9.6865e - 004$. The variance of the restricted estimator, evaluated at the restricted estimates is $9.8080e - 004$. Note that this is in fact higher, where it should be less than or equal. The difference is $-1.2156e - 005$

(i) What is the value of the HW test statistic based on these variances?

The HW statistic is -1.2914 which makes no sense given that the variance is estimated to be negative.

(j) Using the information matrix estimated on the unrestricted estimates to estimate the two variances and calculate the HW statistic based on this.

Now the restricted variance is $9.6863e - 004$ so that the difference is $1.6976e - 008$ (which is positive), leading to a HW statistic of 924.7100. You see how sensitive the HW tests can be. This would not be the case for the other tests (Wald, LR and LM).