**UC Berkeley** **Fall Semester 2004**

**Economics 241A** **Guido Imbens**

# PROBLEM SET I, Part 3

Solutions

1. (a) Suppose that $Y_1, \ldots, Y_N$ are independent and identically distributed random variables, with a distribution with hazard rate $h(t) = \lambda_0 + \lambda_1 \cdot 1\{t > 3\}$, where $1\{\cdot\}$ is the indicator function. Do a likelihood ratio test of the hypothesis that $\lambda_1 = 0$.

   The restricted estimator for $\lambda_0$ is, as we have seen before, $\hat{\lambda}_{0r} = 1/\bar{y}$. Under the alternative the density is

$$f(y; \lambda_0, \lambda_1 = 1\{y \leq 3\}\lambda_0 \exp(-y\lambda_0) + 1\{y > 3\}(\lambda_1 + \lambda_0)\exp(-(y-3)\lambda_1 - y\lambda_0).$$

   The log likelihood function is, letting $N$ denote the sample size, and $N_0$ the number of observations with $y_i \leq 3$,

$$L(\lambda_0, \lambda_1) = \sum_{i:y_i \leq 3} \ln \lambda_0 - y_i\lambda_0 + \sum_{i:y_i>3} \ln(\lambda_1 + \lambda_0) - (y_i - 3)\lambda_1 - y_i\lambda_0$$

$$= N_0 \cdot \ln \lambda_0 - \lambda_0 \sum_i y_i + (N - N_0)\ln(\lambda_1 + \lambda_0) - \lambda_1 \sum_{i:y_i>3}(y_i - 3)$$

   The unrestricted estimators are

$$\hat{\lambda}_{1u} = \frac{N - N_0}{\sum_{i:y_i>3}(y_i - 3)} - \frac{N_0}{\sum_i y_i - \sum_{i:y_i>3}(y_i - 3)},$$

   and

$$\hat{\lambda}_{0u} = \frac{N_0}{\sum_i y_i - \sum_{i:y_i>3}(y_i - 3)}.$$

   Substituting this into the log likelihood function and taking twice the difference gives the likelihood ratio statistic:

   (b) Check the second result in the Appendix to lecture notes 3.

   We want to prove

$$A^{-1} - (A + B)^{-1} = A^{-1}(A^{-1} + B^{-1})^{-1}A^{-1}.$$

   Pre and post multiply by $A$, so we have to prove

$$A - A(A + B)^{-1}A = (A^{-1} + B^{-1})^{-1}.$$

Premultiply by $(A^{-1}+B^{-1})$, so this is equivalent to

$$(A^{-1}+B^{-1})(A-A(A+B)^{-1}A)=I,$$

or

$$I+B^{-1}A-(A+B)^{-1}A-B^{-1}A(A+B)^{-1}A=I,$$

which is equivalent to

$$B^{-1}A-(A+B)^{-1}A-B^{-1}A(A+B)^{-1}A=0.$$

Post multiply by $A^{-1}$ to show that this is equivalent to

$$B^{-1}-(A+B)^{-1}-B^{-1}A(A+B)^{-1}=0.$$

Post multiply by $(A+B)$ to make this equivalent to

$$B^{-1}A+B^{-1}B-I-B^{-1}A=0,$$

which follows directly.

2. Use the same data as in Part 1 of the problem set.

   (a) Discard all observations with durations less than 80. This should leave you with 3338 observations. What is the likelihood function, taking into account the (right) censoring and the (left) truncation, based on an exponential distribution with hazard rate $\lambda$?

   The likelihood function is

   $$\mathcal{L}(\lambda)=\prod_{i=1}^{N}\lambda^{1-c_i}\exp(-(y_i-80)\lambda).$$

   (b) Suppose the hazard function is, as in the last problem set, $\exp(-\beta_0-\beta_1\text{AGE}-\beta_2\text{ED}-\beta_3\text{WHITE}-\beta_4\text{LOCRATE})$. Show that at $(1,0,0,0,0)$ the log likelihood function is -1.9839e+005.

   See matlab output

   (c) Calculate the analytic first derivatives at this point.

   See Table 1.

Table 1: ANALYTICAL AND NUMERICAL FIRST (TIMES $10^{-6}$)

| | analytical | numerical |
|---|---|---|
| intercept | -0.4249 | -0.4249 |
| age | -9.6019 | -9.6017 |
| educ | -5.4887 | -5.4887 |
| white | -0.2275 | -0.2275 |
| locrate | -1.3956 | -1.3956 |

(d) Compare the analytic first derivatives to numerical first derivatives, using $c = 0.000001$.

See Table 1.

(e) Compute the maximum likelihood estimates using the DFP algorithm with starting values $(1, 0, 0, 0, 0)$, and the identity matrix for $A_0$. At each step report $\beta_k$, $p_k$, $d_k$, $q_k$, $\lambda_k$ and $A_k$. Make sure you do the normalization of the direction vector. Use the golden section method to calculate the optimal step length. In the step length calculate stop when the upper and lower bound are less than 0.00001 apart. Stop the main algorithm if the sum of the absolute values of the derivatives is less than 0.001.

See Matlab output for iterations, and Table 2 for final estimates. This took 14 iterations.

Table 2: UNRESTRICTED AND RESTRICTED ESTIMATES

| | unrestricted | | restricted | |
|---|---|---|---|---|
| | est | se | est | se |
| intercept | 4.8210 | (0.1822) | 4.6978 | (0.1583) |
| age | -0.0220 | (0.0082) | -0.0200 | (0.0081) |
| educ | 0.1622 | (0.0127) | 0.1608 | (0.0125) |
| white | -0.0286 | (0.0435) | | |
| locrate | -0.0250 | (0.0187) | | |

(f) Estimate the standard errors for the ml estimates using the average of the second derivatives to estimate the information matrix.

See Table 2.

(g) Now we will consider testing the hypothesis $H_0$: $\beta_3 = \beta_4 = 0$, against the alternative hypothesis that either $\beta_3$ or $\beta_4$ (or both) differs from zero, at the 5% level. First estimate the parameters under the null hypothesis. using the DFP algorithm with starting values $(1, 0, 0)$, and the identity matrix for $A_0$. At each step report $\beta_k$, $p_k$, $d_k$, $q_k$, $\lambda_k$ and $A_k$. Make sure you do the normalization of the direction vector. Use the golden section method to calculate the optimal step length. In the step length calculate stop when the upper and lower bound are less than 0.00001 apart. Stop the main algorithm if the sum of the absolute values of the derivatives is less than 0.001.

This took 10 iterations. See Table 2.

(h) Estimate the standard errors for the restricted ml estimates using the average of the second derivatives to estimate the information matrix.

See Table 2.

(i) Next do the likelihood ratio test.

The unrestricted log likelihood function is -1.5853e+004. The restricted log likelihood function is -1.5855e+004. Twice the difference is $LR = 2.4222$. At the 5% we do not reject the null hypothesis. The critical value of the chi-squared(2) distribution is 5.99.

(j) Do the Wald test, using the second derivatives and estimates under the alternative to estimate the information matrix.

The Wald statistic is $WALD = 2.4352$. Again we do not reject the null hypothesis at the 5% level.

(k) Do the LM or score test, using the $N$ times $R^2$ representation.

The LM statistic is $LM = 1.7163$. Again we do not reject the null hypothesis.