UNIVERSITY OF NORTH CAROLINA
Department of Economics

Economics 275 | Dr. Gallant
Homework 1 | Fall 2000
Due September 13, 2000

Marron's argument that least squares cross validation is too variable for density bandwidth choice may be incorrect in this sense: He seeks to find the best population bandwidth but least squares cross validation actually seeks to find the best bandwidth on a per sample basis. Thus, the variablity that Marron complains of might just be the result of least squares cross validation doing its job properly. Check this out as follows:

Assignment:

Generate plot points $f^o(x_i)$ on an equally spaced fine grid $x_0 < x_1 < \cdots < x_m$ and a sample of size $n = 1000$ from one of the smoother densities in the Marron-Wand test suite using mixture. For a grid of bandwidths $h_0 < h_1 < \cdots < h_\ell$, use uniden to compute both $M_1(h_i)$ and the kernel estimate $\hat{f}_{h_j}(x_i)$ on exactly the same grid $x_0 < x_1 < \cdots < x_m$ as used for mixture above. Putting $\Delta x = x_1 - x_0$, $\mathrm{MSE}(h_j)$ can be computed using

$$\mathrm{MSE}(h_j) \doteq \sum_{i=0}^{m} \left[f^o(x_i) - \hat{f}_{h_j}(x_i) \right]^2 \Delta x.$$

Why? Plot, on the same graph, $\mathrm{MSE}(h_j)$ and

$$\widehat{\mathrm{MSE}}(h_j) = M_1(h_i) + \sum_{i=0}^{m} \left[f^o(x_i) \right]^2 \Delta x.$$

Indicate on the graph the cross validated bandwidth selection, the Sheather-Jones bandwidth selection, and the Silverman bandwidth selection.

Mixture and uniden are available by anonymous ftp from ftp.econ.duke.edu in directory pub/arg/npe or click “Browse ftp site” on the course web page, www.unc.edu/ arg/econ275.