
Reading: Chapters 9 and 11 of Cameron and Trivedi.

Problem-1: Let $Z_1, \ldots, Z_n$ be i.i.d. copies of a random variable $Z \sim f(z)$ where $Z = (Y, X)$.

(a) What is the kernel density estimator of $f(z_0)$?

(b) What is the bias and variance of this estimator?

(c) What is the optimal bandwidth that minimizes the MSE of this estimator? Explain the curse of dimensionality by comparing the rate of convergence of the estimator of $f(z_0)$ using optimal bandwidth with that done in the class for a scalar Z.

(d) We are often interested in finding $f(z)$ for the entire support of Z. Assume that this support is compact. Find out the optimal bandwidth by minimizing the IMSE instead of the MSE.

Problem-2: Problem 9-3 of Cameron and Trivedi (page 335). Use Matlab to answer this question. You may verify your results from Matlab using Stata.

Problem-3: You may find Gregory and Veall (1985, Formulating Wald Tests of Nonlinear Restrictions, *Econometrica*, 53, 1465–1468.) helpful for this problem.

Suppose that you have n observations $\{(y_t, X_{1t}, X_{2t}) : t = 1, \ldots, n\}$ from the linear regression model

$$y_t = \beta_0 + \beta_1 X_{1,t} + \beta_2 X_{2,t} + \epsilon_t. \tag{1}$$

You want to test $H_g : g(\beta_1, \beta_2) = 0$ against $K_g : g(\beta_1, \beta_2) \neq 0$ where $g(\beta_1, \beta_2) = \beta_1 - 1/\beta_2$. Note that the same test can be alternatively be done by testing $H_h : h(\beta_1, \beta_2) = 0$ against $K_h : h(\beta_1, \beta_2) \neq 0$ where $h(\beta_1, \beta_2) = \beta_1\beta_2 - 1$.

Consider the following DGP based on (1):

$$\beta_0 = 1, \beta_1 = 10, \beta_2 = .1,$$
$$\begin{bmatrix} X_{1,t} \\ X_{2,t} \end{bmatrix} = \begin{bmatrix} .6 & .3 \\ .3 & .6 \end{bmatrix} \begin{bmatrix} X_{1,t-1} \\ X_{2,t-1} \end{bmatrix} + \begin{bmatrix} V_{1,t} \\ V_{2,t} \end{bmatrix}$$

where $(\epsilon_t, V_{1,t}, V_{2,t}) \overset{\text{i.i.d.}}{\sim} N(0, I_3)$.

(a) Draw $n = 20$ observations from the model and perform the Wald tests for H_g and H_h. Keeping everything else fixed, repeat the experiment 1000 times by drawing a new series of $\{\epsilon_t\}$ each time. Perform the two Wald tests every time. Report the empirical size of the two tests.

(b) Redo part (a) for each sample size $n \in \{20, 30, 40, \ldots, 400\}$ and draw a figure as follows. The x-axis has a grid for n, while the y-axis has the empirical size. You plot two lines: the empirical size based on the g-type test and that based on the h-type test. Hint: When $n = 400$, the empirical size of both tests should be fairly close to the nominal size, 5%.

(c) Now let us see if bootstrap can help us obtain better approximation of nominal size when n is as small as 20. To do residual-based bootstrap, take the number of bootstrap replication $B = 999$ (this is entirely under your control). Follow the algorithm described below for H_g.[1]

Step 1. Estimate $\beta_0, \beta_1, \beta_2$ imposing H_g and obtain the residuals via nonlinear least squares.[2] Call them $\tilde{\beta}_0$, $\tilde{\beta}_1$, $\tilde{\beta}_2$, and $\{\tilde{\epsilon}_t\}$.

[1] The procedure is slightly different from what was taught in Saraswata's class since the regressors here have time series structures.

[2] Kaiji's comment: In my last recitation I told you to run OLS and then impose H_g, but that has turned out to be wrong. We should impose H_g and then run NLS to get restricted estimates. Sorry for the confusion.

Step 2. For the b-th bootstrap replication, draw a random sample of the residuals: $\widetilde{\epsilon}^*_b = (\widetilde{\epsilon}^*_{b1}, \ldots, \widetilde{\epsilon}^*_{bn})'$. Note that $\widetilde{\epsilon}^*_{bt}$ is uniformly drawn from $\{\tilde{\epsilon}_1, \ldots, \tilde{\epsilon}_n\}$ with replacement for each t. Generate $y^*_b = (y^*_{b1}, \ldots, y^*_{bn})'$ according to $y^*_{bt} = \tilde{\beta}_0 + \tilde{\beta}_1 X_{1,t} + \tilde{\beta}_2 X_{2,t} + \widetilde{\epsilon}^*_{bt}$. A key point here is that we are not reordering the regressors in order to preserve their time series structures.

Step 3. Obtain the b-th bootstrap OLS estimate of $\beta_0, \beta_1, \beta_2$ and the corresponding standard errors to construct the t statistic for H_g.[3] Call it $t^{(b)}$.

Step 4. Repeat Steps 2 and 3 for $b = 1, \ldots, B$ and obtain the bootstrap critical value by sorting $\{|t^{(1)}|, \ldots, |t^{(B)}|\}$ and picking the 95% quantile. Test H_g using your original t statistic from part (a), but the critical value obtained here.

Repeat the Monte Carlo experiment 1000 times to obtain the empirical size of the test based on bootstrap critical value. Now do the same for the null hypothesis H_h. Compare your results with that in part (a).

[3] Recall that the t test is a Wald test with only one restriction.