Christiano
FINC 520, Spring 2008
Homework 6, due Thursday, May 22.

1. Consider the iid Normal stochastic process, $\{x_t\}$, with $Ex_t = \mu$ and $E(x_t - \mu)^2 = \sigma^2$. Let $\mu_l = E(x_t - \mu)^l$ denote the l^{th} moment about the mean. A property of the Normal distribution is that odd-ordered moments are zero, and even-ordered moments satisfy:

$$\mu_{2k} = \frac{\sigma^{2k}}{2^k}\frac{(2k)!}{k!}.$$

Skewness of any distribution is defined as

$$s = \frac{E(x_t - \mu)^3}{\sigma^3}.$$

In the case of a Normal distribution, this is of course zero. Skewness is estimated as follows:

$$\hat{s}_T = \frac{\frac{1}{T}\sum_{t=1}^{T}(x_t - \hat{\mu})^3}{\hat{\sigma}^3},\ \hat{\sigma}^2 = \frac{1}{T}\sum_{t=1}^{T}(x_t - \hat{\mu})^2,\ \hat{\mu} = \frac{1}{T}\sum_{t=1}^{T}x_t.$$

(a) Set this estimator of skewness up as an exactly identified GMM estimator. Show that $\hat{s}_T$ is asymptotically normal with standard deviation,

$$\sqrt{\frac{6}{T}}.$$

(b) Now suppose that the true value of the mean is known. Show that $\hat{s}_T$ is asymptotically normal with standard deviation,

$$\sqrt{\frac{15}{T}}.$$

(Hint: to do each part of this question, you have to identify a different 'GMM stochastic process', $h_t(\theta, w_t)$, having the property, $Eh_t(\theta^0, w_t) = 0$, where θ^0 is the true value of the parameters, and $\theta = (\mu, \sigma, s)'$ in the case of part a while $\theta = (\sigma, s)'$ in the case of part b. When computing the matrices required by GMM, be sure to impose all the properties of the Normal distribution. Also, when inverting matrices, make life simple for yourself by using the 'evaluate' command in the 'compute' pull-down menu in Scientific Word.)

(c) Does the sampling distribution of the estimator of σ depend on whether or not μ is known?

2. Consider the data from homework 2. Compute the skewness statistic for each of the four variables (inflation, the long rate, the short rate and GDP growth). To do this, you cannot rely on the assumptions of iidNormality used in the previous question. So, you must take a stand on how many lags to use in constructing the zero-frequency spectral density for the GMM error. For this, use the covariance, and the first, second and third lagged-covariances. Using the GMM sampling theory, compute the $p-$value of the empirical $\hat{s}$ statistic under the null hypothesis that the true value is zero and the sampling distribution of $\sqrt{T} \times \hat{s}$ is $N\left(0, \hat{V}\right)$, where $\hat{V}$ is the GMM estimator of the sampling variance of $\hat{s}$ (*i.e.*, the $p-$value of $\hat{s}$ is the probability that an $N\left(0, \hat{V}/T\right)$ random variable is larger than $\hat{s}$.

3. For question 2, you computed the $p-$value of $\hat{s}$ under the null hypothesis of no skewness, using asymptotic sampling theory. It is possible that T is not large enough for the asymptotic theory to be a good approximation. You can use a Monte Carlo sampling experiment to check this. The idea is to use a computer to directly compute the sampling distribution, in an empirically relevant sample size and under the null hypothesis of no skewness, of the skewness statistic. To do this, one needs to take a position on the mechanism that generated the data used to compute the skewness statistic. This step is important because, presumably, the sampling distribution of the skewness statistic is sensitive to what mechanism generates the data. A natural choice for the data generating mechanism is the four variable VAR(4) you estimated for homework 2:

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \phi_3 y_{t-3} + \phi_4 y_{t-4} + \varepsilon_t,\ E\varepsilon_t\varepsilon_t' = \Omega,$$

where Ω is estimated by

$$\frac{1}{T}Y'MY,$$
$$M \equiv I - X\left(X'X\right)^{-1}X',$$

and y_t, ε_t, X and Y are defined in homework 2. Write

$$CC' = \Omega,$$

where C is lower triangular with positive elements on its diagonal.[1] Note that if $\tilde{\varepsilon}_t$ is $N(0, I)$, then $C\tilde{\varepsilon}_t$ has variance-covariance matrix, Ω. Generate N artificial observations, $y_1, y_2, ..., y_N$, using

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \phi_3 y_{t-3} + \phi_4 y_{t-4} + C\tilde{\varepsilon}_t,$$

where $\tilde{\varepsilon}_t$ is drawn independently from $N(0, I)$ for each t.[2] To start this simulation, you will need initial conditions, $y_0, ..., y_{-3}$. Set these to the value of Ey_t implied by the estimated VAR. Let $N = 200 + T$, where T is the number of rows in X, the matrix containing the actual data. Compute $\hat{s}$ for each of the four variables using $y_{201}, ..., y_{200+T}$. (Note that by beginning with y_{201}, the initial y_t's are in effect being drawn from the unconditional distribution of y_t.) Repeat this exercise 2,000 times. Obtain the $p-$value for each of your four skewness statistics, by computing the fraction of times that the simulated skewness statistics exceed their empirical counterparts. Note that the simulations impose the zero skewness statistic of the null hypothesis because of the way the VAR disturbances are drawn. Do you get a different result from what you found in question 2?

4. Develop asymptotic sampling theory for the centered R^2 statistic on page 202 of Hamilton. Use the asymptotic sampling theory for GMM for this.

---

[1] This matrix exists and is unique, given that Ω is positive definite. The routine, chol(B) in MATLAB computes the upper triangular Choleski decomposition. To obtain what we want, set C =chol(B')'.

[2] You can draw $\tilde{\varepsilon}_t$ by executing the MATLAB command, randn(4,1).

5. Consider the VAR(4) you studied in question 2, part (c), of homework #2. Compute the fitted residuals, $\hat{\varepsilon} = MY$, where $M = I - X(X'X)^{-1}X'Y$. Note that the first column of $\hat{\varepsilon}$ corresponds to the 201 quarterly observations, 1955Q4-2005Q4 on residuals in the short term interest rate equation (i.e., R);the second column corresponds to observations over the same period for fitted residuals in the R^l equation; the third column corresponds to GDP growth and the fourth, to inflation. We will examine the hypothesis that the residuals in the underlying 'true' VAR are identically distributed over time.

   (a) For each of the four residuals, graph their 2-year moving, centered standard deviation, s_t:

   $$s_t = std\left([\hat{\varepsilon}_{t-4}, ..., \hat{\varepsilon}_{t+4}]\right),$$

   for $t = 1956Q4 - 2004Q4$ ('std' in the above expression is the name of the MATLAB command for computing the standard deviation). Does this graph appear to be consistent with the proposition that the underlying disturbances are drawn, at each date, from the same distribution (in particular, that they have the same variance at each date)?

   (b) Implement the Engle test for ARCH described on page 664, with $m = 4$ and $m = 6$, for each of the four fitted residual series. Compute Engle's proposed test statistic on each of the four residual series (i.e., T times the centered R^2 statistic defined in equation [8.1.14], page 202). Compute the $p-$value of each test statistic using the asymptotic χ^2 distribution theory for the null hypothesis that the underlying ε_t's are iid (hint: this is implemented as $1 - chi2cdf(statistic, m)$, in MATLAB). You should ponder (though it need not be worked out for this homework) how you yourself would derive Engle's χ^2 asymptotic sampling theory using our GMM results).

   (c) The asymptotic sampling theory you used in part b of this question is asymptotic, and you may wonder whether you have enough observations for this theory to be a good approximation. There is reason to be concerned because I think that in part b you will find that there is no significant departure from the iid assumption in the GDP growth shock. But, this conclusion seems inconsistent with the results in Justiniano and Primiceri, who argue that the 'Great Moderation' (a fall in the volatility of macroeconomic variables after 1980) reflects a fall in the variance of shocks to GDP post 1980. One possibility worth exploring is that the failure to find significant departures from the iid assumption in the GDP shocks reflects the failure of asymptotic theory (working against this possibility is that I think you'll find that there is not a very large amount of visual evidence against the iid assumption for the GDP shock in the figures you constructed for part a). A Monte Carlo experiment needs to be designed in which an empirically plausible data generating mechanism is used to simulate the Engle statistics under the null hypothesis of iid disturbances (Engle worked this out by hand - using our GMM sampling theory - for the case T large, but we now want to work it out for T of normal size and this is too hard to work out by hand). For this, you should use the bootstrap and a computer.

Doing the bootstrap requires executing the simulations you did for question 3, but with a twist. Instead of drawing the ε_t's from the multivariate Normal distribution, you will now draw them by bootstrap (the Normal distribution made sense in question 3, because the null hypothesis there required that the shocks exhibit no skewness and lack of skewness is not guaranteed in the bootstrap procedure...though for present purposes the bootstrap is just fine because it imposes the iid assumption by design).

The ε_t's can be drawn by bootstrap as follows. Let $\hat{\varepsilon}_1, ..., \hat{\varepsilon}_T$ denote the fitted residuals from the VAR (i.e., $\hat{\varepsilon}_t$ is the t^{th} row of MY). The idea is to draw 2,000 random simulations on the Y_t's, each of length $200 + T$, by simulating the VAR with the estimated parameter values. In each of the 2,000 simulations, you will need $200 + T$ independent draws, with replacement, from the set of vectors, $\hat{\varepsilon}_1, ..., \hat{\varepsilon}_T$. This in effect requires that you draw randomly from the set of integers, $\{1, ..., T\}$. You can do so as follows.

A single random draw from $\hat{\varepsilon}_1, ..., \hat{\varepsilon}_T$ can be accomplish by drawing a uniform (0,1) random variable from Matlab using the MATLAB routine, rand. Multiply this random variable by T. Then, apply the MATLAB routine, ceil, to round the result up to the next highest integer. This gives you one random draw from $\{1, ..., T\}$. Call the result, $j \in \{1, ..., T\}$. Then, use $\hat{\varepsilon}_j$. Repeat this $200 + T$ times, and you have $200 + T$ random draws, with replacement, from $\hat{\varepsilon}_1, ..., \hat{\varepsilon}_T$. These can be used to simulate one realization out of the 2,000 you need.

In each artificial realization, use $Y_{200+1}, ..., Y_{200+T}$ to fit a VAR and obtain VAR residuals. Then, compute the Engle $T \times R^2$ statistic on each residual that you did for part b of this question. The end result will be 2,000 realizations of this Engle statistic for each of the four residuals. Compute $p-$values by calculating the fraction times that the realizations of these statistics exceed the empirical values of the statistics that you computed in part b. If the $p-$values are essentially the same as the ones in part b, then the asymptotic theory is working well (if you did the bootstrap with $T = 1,000,000$, then the asymptotic theory should work perfectly. You may want to check this.)