Harald Schmidbauer İstanbul Bilgi Üniversitesi

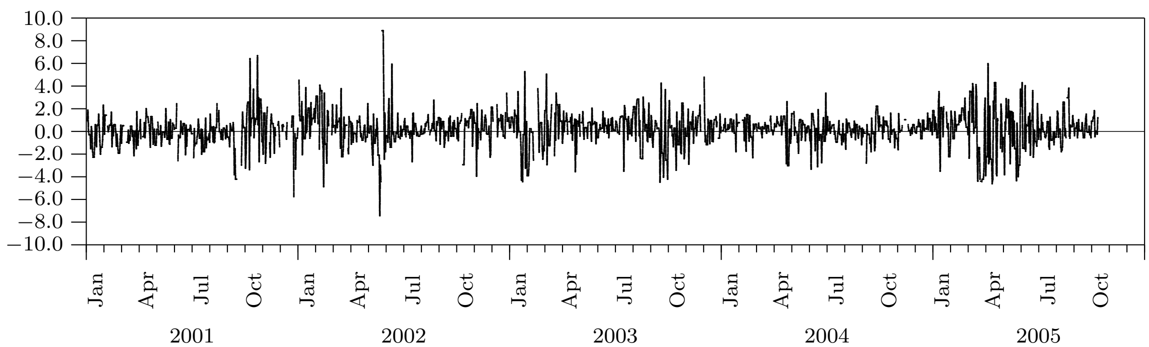
FEC 522: Financial Econometrics II

Spring 2011

Problem Sheet With Solutions

(to be discussed Monday, March 28, 2011)

Problem 1: Describe the properties of the series displayed below in terms of trend, seasonality, stationarity, homoskedasticity. (Please itemize your answer, writing short sentences explaining the series' properties briefly. — This series is the series of daily returns on the stock index KSE, that is: Karachi stock exchange, Pakistan.)



Solution: This question is about the "stylized facts" of the series.

- There is no trend.
- There is no obvious seasonality. (Depending on the goal of our analysis, we might search evidence for intra-week seasonality, which may be found in a series of daily returns of a stock index.)
- The series may be stationary insofar as there does not seem to be a trend in volatility. (This consideration is important when building a suitable stochastic model for the observed series.)
- This series is not homoskedastic.

Problem 2: The following R code simulates a stochastic process:

```
eps = rnorm(500)
x = 0
for (t in 2:500) { x[t] = 0.7 * x[t-1] + eps[t] }
```

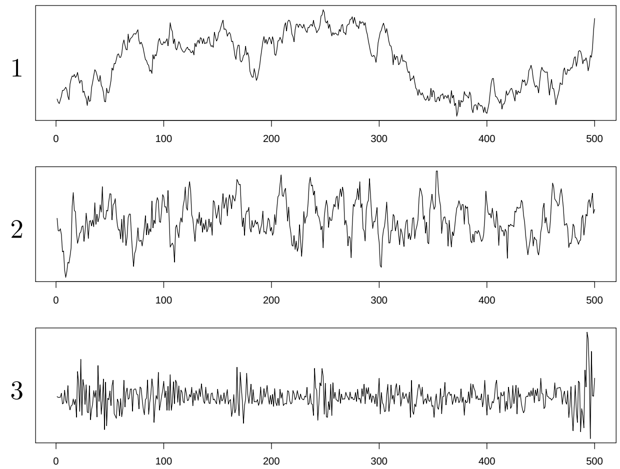
Explain which process is simulated. (Write down the equations which define this process.)

Solution: This code simulates an AR(1) process starting in 0. The model equation is:

$$x_1 = 0, \quad x_t = 0.7 \cdot x_{t-1} + \epsilon_t, \quad t \geq 2,$$

where (ϵ_t) is Gaussian white noise with unit variance.

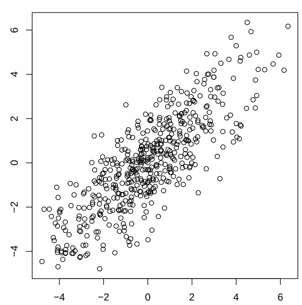
Problem 3: The following three figures show simulated sample paths of stochastic processes:



Which of these processes could be a simulated AR(1) process?

Solution: Process 3 is heteroskedastic, hence cannot be an AR(1) process. Process 1 might not be mean-reverting, so it could only be the sample path of an AR(1) with coefficient close to 1. The most straightfoward candidate for an AR(1) sample path is Process 2 (it look mean-reverting and homoskedastic). Fitting an AR(1), followed by a diagnostic check, would be in order.

Problem 4: The following figure shows the scatterplot of a simulated process (X_t), with X_{t-2} on the x-axis (the abscissa), and X_t on the y-axis (the ordinate). The correlation of the points in the scatterplot is about 0.80.



a) Which stochastic model might have been used to simulate the series? Give reasons for your answer. (Hint: The process is among the following: $X_t = \epsilon_t$ (white noise); $X_t = \epsilon_t + \beta\epsilon_{t-1}$; $X_t = c + aX_{t-1} + \epsilon_t$; $X_t = c + X_{t-1} + \epsilon_t$; $X_t = c - X_{t-1} + \epsilon_t$.)

b) Write down the model equation, using the correlation in the scatterplot and your knowledge of the autocorrelation function.

c) How can this process be simulated practically? (Explain by, for example, writing down a computer program.)

Solution:

a) There is a correlation of 0.8 between X_{t-2} and X_t, i.e. at lag 2. Going through the models given in the hint:

  - $X_t = \epsilon_t$ — impossible; there would be zero correlation at all lags $s \geq 1$. (This is simply an iid sequence.)
  - $X_t = \epsilon_t + \beta\epsilon_{t-1}$ — impossible; there would be zero correlation at all lags $s \geq 2$. (This is an MA(2) process.)

- $X_t = c + aX_{t-1} + \epsilon_t$ — possible, with $a = \sqrt{0.8}$. This gives the observed correlation. (This is an AR(1) process.)
- $X_t = c + X_{t-1} + \epsilon_t$ — the correlation would be closer to 1. (This is a random walk with drift.)
- $X_t = c - X_{t-1} + \epsilon_t$ — the correlation would be negative.

b) These considerations lead us to the model $X_t = c + \sqrt{0.8}X_{t-1} + \epsilon_t$.

c) Simulation program: Cf. Problem 2 (with the coefficient suitably modified, of course.)

Problem 5: Let a stochastic process (X_t) be given as follows: $X_t = 3 + 0.5X_{t-1} + \epsilon_t$, where (ϵ_t) is white noise with $\text{var}(\epsilon_t) = 1$.

a) What process is (X_t)?

b) Explain briefly why this is a "conditional expectation" model, but not a "conditional variance" model.

c) Suppose we observed $X_t = 2.5$, $X_{t-1} = 1.7$. Compute a forecast for X_{t+1}.

d) Suppose we observed $X_t = 2.5$, $X_{t-1} = 1.7$. Compute a forecast for the variance of the process at time $t + 1$, that is, for the variance of X_{t+1}.

Solution:

a) This is an AR(1) process.

b) This is a conditional expectation model, because

$$\text{E}(X_{t+1}|X_t, X_{t-1}, X_{t-2}, \dots) = 3 + 0.5X_t,$$

which does use past information (namely X_t). It is not a conditional variance model, because

$$\text{var}(X_{t+1}|X_t, X_{t-1}, X_{t-2}, \dots) = \text{var}(\epsilon_{t+1}),$$

independent of the past.

c) The forecast is: $\text{E}(X_{t+1}|X_t = 2.5, X_{t-1} = 1.7) = 3 + 0.5 \cdot 2.5$.

d) The variance forecast is 1.

Problem 6: Which of the following processes $(X_t)_t$ is weakly stationary?

A: $X_t = 1.6 + X_{t-1} + \nu_t$

B: $X_t = 0.6 + X_{t-1} + \nu_t$

C: $X_t = 0.8X_{t-1} + \nu_t$

D: $X_t = 0.8\nu_t + 0.6\nu_{t-1}$

E: $X_t = \nu_t\sqrt{2 + 2X_{t-1}^2}$

The term (ν_t) is always assumed to be white noise with variance one.

Solution: A, B: no (both are random walks with drift); C: yes (AR(1)); D: yes (MA(1)); E: no (equation points to ARCH(1), but non-stationary since $\alpha_1 = 2 > 1$).

Problem 7: Consider the model $X_t = \nu_t \cdot \sqrt{\alpha_0 + \alpha_1 X_{t-1}^2}$, where (ν_t) is Gaussian white noise with $\text{var}(\nu_t) = 1$.

a) Which model is this?

b) If the constants $\alpha_0, \alpha_1 \in \mathbb{R}$ are appropriately chosen, can we hope to have a good model for a heteroskedastic series?

Solution:

a) This is an ARCH(1) model.

b) Yes, this model displays heteroskedasticity in the form that high variability will persist for some time. However, it may be inferior to a GARCH model in terms of AIC and its capability to extract autocorrelation from the squared residual series.

Problem 8: A GARCH model was fitted to the series of daily returns in percent on a stock. The estimation output, assuming Gaussian white noise, is:

```
    Estimate  Std. Error  t value Pr(>|t|)
a0  0.027832    0.006282     4.43 9.41e-06 ***
a1  0.151097    0.009064    16.67  < 2e-16 ***
b1  0.830477    0.008813    94.24  < 2e-16 ***
```

The mean return is 0.017.

a) Write down the model equations defining the process for the return series.

b) According to this model: What is the long-run variance of returns?

c) Suppose today's return was 1.5%; the variance forecast for today was $h_t = 1.1619$. Determine tomorrow's return distribution.

d) The next day's return was actually +4.6%. What is your conclusion with respect to the distribution you found in (c)?

(Hints: $\epsilon_t = \nu_t \cdot \sqrt{h_t}$, $h_t = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 h_{t-1}$; $E(\epsilon_t) = 0$; $E(\epsilon_t|\epsilon_{t-1}) = 0$; $\text{var}(\epsilon_t) = \frac{\alpha_0}{1-\alpha_1-\beta_1}$; $\text{var}(\epsilon_t|\epsilon_{t-1}) = h_{t-1}$.)

Solution:

a) Model equations:

$$X_t = 0.017 + \epsilon_t, \quad \epsilon_t = \nu_t \cdot \sqrt{h_t}, \quad h_t = 0.0278 + 0.1511 \cdot (X_{t-1} - 0.017)^2 + 0.8305 \cdot h_{t-1},$$

where (ν_t) is white noise with $\text{var}(\nu_t) = 1$. (X_t designates the return on day t in percent.)

b) The long-run variance of returns is given by

$$\text{var}(\epsilon_t) = \frac{\alpha_0}{1 - \alpha_1 - \beta_1} = \frac{0.0278}{1 - 0.1511 - 0.8305} = 1.51.$$

c) Tomorrow's variance forecast is:

$$h_{t+1} = 0.0278 + 0.1511 \cdot (1.5 - 0.017)^2 + 0.8305 \cdot 1.1619 = 1.3251,$$

so that $X_{t+1} \sim \text{N}(0.017, 1.3251)$.

d) The return of +4.6% is outside the three-sigma interval for $\text{N}(0.017, 1.3251)$, thus having a very small probability. In other words, a return of +4.6% is not what we believed would likely happen. We would try to identify an event (political, for example) which made the stock price jump by 4.6%.

Problem 9: The following R code simulates a GARCH process:

```
N = 1000
nu = rnorm(N + 50)
eps = rep(0,N + 50)
h = rep(0, N + 50)
a0 = 1
a1 = 0.1
a2 = 0.07
b1 = 0.82

for (i in 3:(N + 50))  {
  h[i] = a0 + a1*eps[i-1]^2 + a2*eps[i-2]^2 + b1*h[i-1]
  eps[i] = nu[i]*sqrt(h[i])
}

eps = eps[-(1:50)]
```

a) What is the order of this process?

b) Fit a GARCH model to this process. Check the series of squared residuals for autocorrelation.

c) Fit a GARCH model of (wrong!) order (1,1) to the simulated series. (This model is mis-specified, since we know in this case how we created the series.) Check the series of squared residuals for autocorrelation again. Do we see the model is mis-specified?

d) Now increase N to 10000 and do (c) again. (The power of the test of H_0: "There is no autocorrelation in the series of squared residuals" increases with the length of the series.)

Solution: Available online.
(http://www.hs-stat.com/courses/FEC522/; click on "sol_p9_problem_sheet_2011-03-26.R".)

Problem 10: File `data_dji_fchi_ftse_gdaxi_hsi_ssec_sti_daily.csv` contains data from stock indices DJIA (dji), CAC 40 (fchi), FTSE 100 (ftse), DAX (gdaxi), HSI (hsi), SSEC (ssec), and STI (sti). Fit a suitable GARCH model to each return series.

Solution: Extract a return series from file `data_dji_fchi_ftse_gdaxi_hsi_ssec_sti_daily.csv` by something like

```
my.data = read.table('data_dji_fchi_ftse_gdaxi_hsi_ssec_sti_daily.csv', header = T, sep = ',')
my.return.series = my.data$dji.ret
```

Then mean-correct the series:

```
x = my.return.series - mean(my.return.series)
```

The series `x` can then be analyzed as outlined in the solution to Problem 9.