FM 431: Econometrics of Financial Markets

Fall 2009

PROBLEM SHEET # 10

Problem 1: The following model was fitted to a series (1000 observations) of daily returns in percent on an asset, X_t:

$$X_t = 0.1 + \epsilon_t, \quad \epsilon_t = \nu_t \cdot \sqrt{h_t}, \quad h_t = 2.1 + 0.09\epsilon_{t-1}^2 + 0.89 h_{t-1},$$

where (ν_t) is Gaussian white noise.
For day $t = 1000$, a return of $x_{1000} = -1.3$ was observed, and h_{1000} was estimated as 5.

a) Determine the distribution of X_{1001}.

b) Now suppose that the actually observed return was $x_{1001} = 2.5\%$. Should we be surprised? Give reasons for your answer.

c) Using this example, explain why GARCH is a conditional variance model.

d) Which information do the model equations give us about the *unconditional* variance of daily returns?

e) Could this model be used to make a statement about the variance of the daily return one year (250 days) ahead? Why, why not?

Problem 2: A seasonal ARIMA was fitted to a time series of monthly marriages in Turkey, starting with January 1998 and ending with December 2007. (The logarithm of monthly values was taken before fitting the model.) A reasonably good model has order $(1,1,1) \times (1,0,0)_{12}$; the corresponding R output is:

```
Call:
arima(x = log(my.series), order = arima.order, seasonal = arima.seasonal)

Coefficients:
          ar1      ma1    sar1
      -0.2254  -0.7868  0.9152
s.e.   0.1223   0.0816  0.0295

sigma^2 estimated as 0.02542:  log likelihood = 38.13,  aic = -68.25
```

a) Write down the model equation for the process (X_t), $X_t = \ln$ of the number of marriages in month t. (Use the lag operator L.)

b) Would you expect autocorrelation to be present in the residuals? Why / why not? How can you test for autocorrelation?

c) Is the process (X_t) stationary? Give reasons for your answer.

d) With this model, a forecast for $X_{\text{Jan 2010}}$ is 10.588, with a standard error (estimated standard deviation) of 0.358. Using these values, obtain a point forecast for the number of marriages in January 2010. What can we do with the standard error?

Problem 3: Three regression models were fitted to weekly data from January 2006 through September 2009 (190 weeks):

Model 1: Regression of WTI crude oil prices w.r.t. Brent crude oil prices:

```
Call:
lm(formula = wti.level ~ brent.level)

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -1.135955   0.684564  -1.659   0.0987 .
brent.level  1.025863   0.008832 116.159   <2e-16 ***
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

Residual standard error: 2.786 on 188 degrees of freedom
Multiple R-squared: 0.9863,Adjusted R-squared: 0.9862
F-statistic: 1.349e+04 on 1 and 188 DF,  p-value: < 2.2e-16
```

Model 2: Regression of WTI crude oil prices w.r.t. gold prices:

```
Call:
lm(formula = wti.level ~ gold.level)

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 30.62729    9.23917   3.315   0.0011 **
gold.level   0.05792    0.01192   4.861 2.46e-06 ***
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

Residual standard error: 22.4 on 188 degrees of freedom
Multiple R-squared: 0.1116,Adjusted R-squared: 0.1069
F-statistic: 23.63 on 1 and 188 DF,  p-value: 2.459e-06
```

Model 3: Regression of WTI crude oil price returns w.r.t. gold price returns:

```
Call:
lm(formula = wti.ret ~ gold.ret)

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.26601    0.48733   0.546    0.586
gold.ret    -0.01103    0.14702  -0.075    0.940

Residual standard error: 6.674 on 188 degrees of freedom
Multiple R-squared: 2.996e-05,Adjusted R-squared: -0.005289
F-statistic: 0.005633 on 1 and 188 DF,  p-value: 0.9403
```

a) Which of these three models would you expect to be meaningful, which meaningless? Give reasons for your answer. (Hint: Use the concept of cointegration in your discussion.)

b) Which further test (or class of tests) would you use to confirm your statements in (a)? Explain the null hypothesis, the alternative hypothesis, and the data you use to carry out the test.

c) In Model 3: What does the number 0.940 in the `Pr(>|t|)` column of the output mean?