Time series & their characteristics

A **time series** consists of repeated observations on the same entity collected over a number of periods.

Economic time series usually exhibit some form of serial correlation. Measures of such correlation are called **serial** or **autocorrelation** coefficients.

The correlation coefficient between observations on two variables $x_1, \ldots, x_n$ and $y_1, \ldots, y_n$ is given by

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\left(\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2\right)^{\frac{1}{2}}}.$$

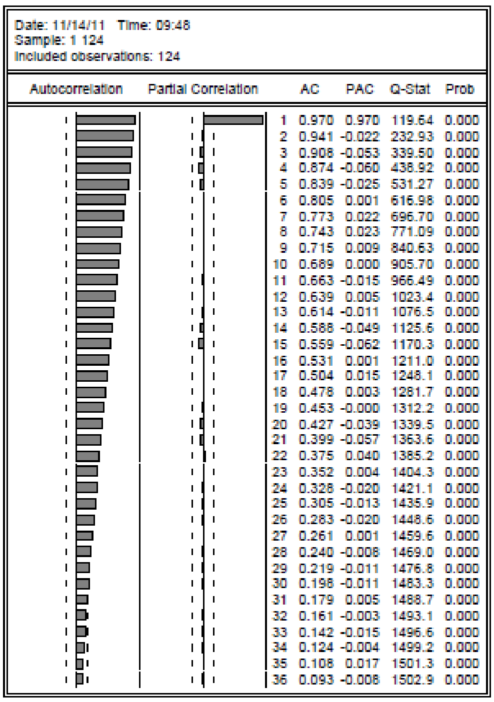
Adapting this to the case of the correlation between y and its own past (lagged) value i.e. between $y_2, \ldots, y_T$ and $y_1, \ldots, y_{T-1}$ produces

$$r = \frac{\sum (y_t - \overline{y_t})(y_{t-1} - \overline{y_{t-1}})}{\left(\sum (y_t - \overline{y_t})^2 \sum (y_{t-1} - \overline{y_{t-1}})^2\right)^{\frac{1}{2}}}.$$

This is the correlation between y_t and y_{t-1}, and we can form the correlation between y_t and y_{t-2}, between y_t and y_{t-3}, etc. similarly.

Together these constitute the serial or autocorrelations of order 1, 2, 3, etc. The plot of the serial correlations against their order is called the **correlogram**. EViews provides such plots under Quick/Series Statistics/Correlogram

Correlogram of UK


Date: 11/14/11 Time: 09:48
Sample: 1 124
Included observations: 124

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.970	0.970	119.64	0.000
		2	0.941	-0.022	232.93	0.000
		3	0.908	-0.053	339.50	0.000
		4	0.874	-0.060	438.92	0.000
		5	0.839	-0.025	531.27	0.000
		6	0.805	0.001	616.98	0.000
		7	0.773	0.022	696.70	0.000
		8	0.743	0.023	771.09	0.000
		9	0.715	0.009	840.63	0.000
		10	0.689	0.000	905.70	0.000
		11	0.663	-0.015	966.49	0.000
		12	0.639	0.005	1023.4	0.000
		13	0.614	-0.011	1076.5	0.000
		14	0.588	-0.049	1125.6	0.000
		15	0.559	-0.062	1170.3	0.000
		16	0.531	0.001	1211.0	0.000
		17	0.504	0.015	1248.1	0.000
		18	0.478	0.003	1281.7	0.000
		19	0.453	-0.000	1312.2	0.000
		20	0.427	-0.039	1339.5	0.000
		21	0.399	-0.057	1363.6	0.000
		22	0.375	0.040	1385.2	0.000
		23	0.352	0.004	1404.3	0.000
		24	0.328	-0.020	1421.1	0.000
		25	0.305	-0.013	1435.9	0.000
		26	0.283	-0.020	1448.6	0.000
		27	0.261	0.001	1459.6	0.000
		28	0.240	-0.008	1469.0	0.000
		29	0.219	-0.011	1476.8	0.000
		30	0.198	-0.011	1483.3	0.000
		31	0.179	0.005	1488.7	0.000
		32	0.161	-0.003	1493.1	0.000
		33	0.142	-0.015	1496.6	0.000
		34	0.124	-0.004	1499.2	0.000
		35	0.108	0.017	1501.3	0.000
		36	0.093	-0.008	1502.9	0.000

The autocorrelation column shows the

correlogram for a series on UK income per head 1871-1993.

The autocorrelations show a steady decline. The **partial** correlation entry for 2 (say) shows the correlation between y_t and y_{t-2} when the other lagged $y's$ have been taken into account or partialled out. The near 0 values for all the partials beyond 1 suggest that y_t and y_{t-2} (say) are only correlated because each is correlated with y_{t-1}. When that correlation is taken into account by partialling out the correlation collapses.

There are numerous models for univariate time series. One of the most widely used is the (1st order) **autoregressive** process (met earlier)

$$y_t = \rho y_{t-1} + \varepsilon_t$$

where the $\varepsilon_t's$ are serially independent. An AR1 produces the pattern in the correlogram for UK income per head.

Consider how y_t and y_{t-2} are connected in the autoregressive model:

$$\begin{aligned} y_t &= \rho y_{t-1} + \varepsilon_t \\ y_{t-1} &= \rho y_{t-2} + \varepsilon_{t-1} \end{aligned}$$

$$\Rightarrow y_t = \rho^2 y_{t-2} + \rho\varepsilon_{t-1} + \varepsilon_t.$$

If y_{t-2} is known that would help in forecasting y_t–in the absence of other information. But the variables y_t and y_{t-2} are correlated because of the link through y_{t-1} and apart from this link the variables are unrelated. The partial correlation between them (partialling out y_{t-1}) is zero. If y_{t-1} is known, y_{t-2} would **not** help in forecasting y_t.

Extending the back substitution procedure

$$y_t = \rho^2 y_{t-2} + \varepsilon_t + \rho\varepsilon_{t-1}$$

$$\Rightarrow y_t = \rho^t y_0 + \varepsilon_t + \rho\varepsilon_{t-1} + \ldots + \rho^{t-1}\varepsilon_1$$

shows how the current value y_t is built up from the starting or initial value y_0 and all the shocks $\varepsilon_1, \ldots, \varepsilon_t$ that have hit the system since.

This representation can also be used to explain the pattern in the serial correlations. Examining

$$y_t = \rho^t y_0 + \varepsilon_t + \rho\varepsilon_{t-1} + \ldots + \rho^{t-1}\varepsilon_1$$

$$y_{t-s} = \rho^{t-s} y_0 + \varepsilon_{t-s} + \rho\varepsilon_{t-s-1} + \ldots + \rho^{t-s-1}\varepsilon_1$$

shows that both y_t and y_{t-s} are built from $\varepsilon_1, \ldots, \varepsilon_{t-s}$ although y_t is built from later shocks too. However large s is there is some correlation though for large s there are few common shocks.