Exercises - Time series analysis

DESCRIPTIVE ANALYSIS OF A TIME SERIES

(1) Estimate the trend of the series of gasoline consumption in Spain using a straight line in the period from 1945 to 1995 and generate forecasts for 24 months. Compare the results with Holt's method.

(2) Apply a decomposition method to the gasoline consumption series and estimate the seasonal coefficients. Compare the results when Holt's method and using moving averages.

(3) Obtain the periodogram for the series of Columbus's voyage and interpret it.

(4) Obtain the periodogram for the gasoline consumption series and interpret it.

(5) Prove that the predictions made using Holt's method verify the recursive equation- $\widehat{z}_{T+1}(1) = \widehat{z}_T(1) + \widehat{\beta}_T + \gamma(1-\theta)(z_{T+1-}\widehat{\mu}_T - \widehat{\beta}_T)$.

Exercises - Time series analysis

(1) Figure 1 shows monthly rainfall in Santiago de Compostela during the years 1988-1997. Would it be a stationary process?

(2) Let us consider the process $z_t = .5z_{t-1} + a_t$ where a_t is a white noise process. Calculate the marginal mean, its variance and first order autocovariance. Is the process stationary?

(3) In the above exercise, calculate the expectation and variance of the conditional distribution $f(z_t|z_{t-1})$. Compare these results with those obtained in the above exercise for the marginal distribution.

(4) Let us consider the process $z_t = -.5a_{t-1} + a_t$. Calculate its mean and first and second order autocovariance. Is the process stationary?

(5) Prove that the above process can be written as $z_t = -.5(z_{t-1} + .5a_{t-2}) + a_t = -.5z_{t-1} - .25a_{t-2} + a_t$. Use this equation to calculate the expectation and variance of the conditional distribution $f(z_t|z_{t-1})$. Compare the results with those of the marginal distribution.

(6) Given a stationary series, prove that if the autocovariances are all positive then the mean of the process will be estimated with greater variance than if all the autocovariances are null.

(7) Prove that the mean of the process in exercise (4) is estimated with less variability than in a process with independent data and the same marginal variance.

FIGURE 1. Monthly rainfall in Santiago de Compostela during 1988-1997
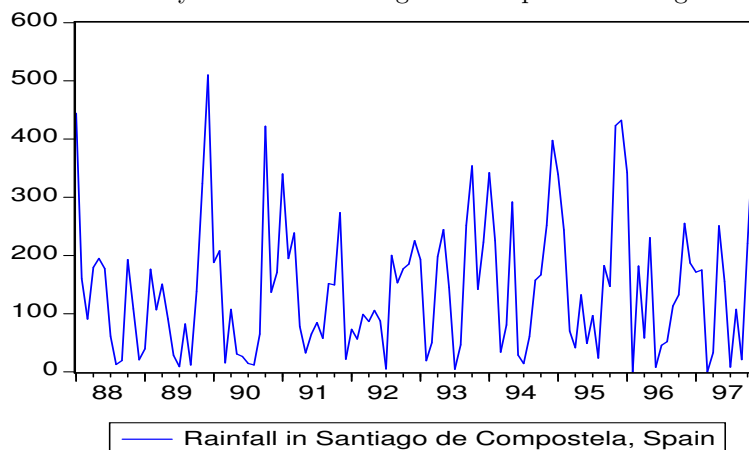
Exercises - Time series analysis

Autoregressive processes

(1) Generate samples of an AR(1) process with $\phi = .7$ using a computer program as follows: (1) generate a vector a_t of 150 random normal variables (0,1); (2) take $z_1 = a_1$; (3) for $t = 2, ..., 150$ calculate $z_t = .7z_{t-1} + a_t$; (4) to avoid the effect of the initial conditions, eliminate the first 50 observations and take the values $z_{51}, ... z_{150}$ as a sample of 100 of the AR(1) process.

(2) Obtain the theoretical autocorrelation function of the process $z_t = .7z_{t-1} + a_t$, where a_t is white noise. Compare the theoretical results with those observed in the sample from the above exercise.

(3) Obtain the theoretical autocorrelation function of the process $z_t = .9z_{t-1} - .18z_{t-1} + a_t$, where a_t is white noise. Generate a realization of the process using a computer and compare the sample function with the theoretical one.

(4) Express in operator notation the processes of exercises (2) and (3). Represent the process in the form $z_t = \sum \psi_i a_{t-i}$ obtaining the inverse operators.

(5) Write the theoretical autocorrelation function of the process $(1 - 1.2B + .32B^2)z_t = a_t$. Obtain the representation of this process as $z_t = \sum \psi_i a_{t-i}$ and comment on the relationship between the coefficients ψ and the autocorrelation function.

(6) Prove that the process $y_t = z_t - z_{t-1}$, where $z_t = .9z_{t-1} + a_t$ is stationary, remains stationary.

(7) Prove that the above process can be written as $y_t = -.09z_{t-2} + a_t - .1a_{t-1}$.

(8) Calculate the inverse operator of $(1 - .8B)(1 - B)$.

(9) Justify whether the process $(1 - .5B)(1 - .7B)(1 - .2B)z_t = a_t$ is stationary and write it in its usual expression.

(10) Calculate the theoretical coefficients of partial autocorrelation for the following AR(2) process: $z_t = .7z_{t-1} - .5z_{t-1} + a_t$, where a_t is white noise.

(11) Prove that if a process is AR(1) if we then make the regression $z_t = \beta z_{t+1} + u_t$ we obtain $\widehat{\beta} = \phi$ and $\text{var}(u_t) = \gamma_0(1 - \phi^2)$, where γ_0 is the variance of the process.

(12) Prove that if a process is AR(1) and we make the regression $z_t = \alpha z_{t-1} + \beta z_{t+1} + u_t$, we obtain $\widehat{\beta} = \widehat{\alpha} = \phi/(1 + \phi^2)$ and $var(u_t) = \gamma_0(1 - \phi^2)/(1 + \phi^2)$. Observe that the variance of the innovations is now less than that of an AR(1) and less than in above exercise and interpret this result.

4

Exercises - Time series analysis

Moving average and ARMA processes

(1) Given the process of zero mean $z_t = (1 - .7B)a_t$: (a) calculate the autocorrelation function; (b) write it as an AR(∞) process.

(2) Prove that the MA(1) processes $z_t = a_t - .5a_{t-1}$ and $z_t = a_t - 2a_{t-1}$ have the same autocorrelation structure but that one is invertible and the other is not.

(3) Prove that the two processes $z_t = a_t + .5a_{t-1}$ and $z_t = .5a_t + a_{t-1}$ are indistinguishable since they have the same variance and the same autocorrelation structure.

(4) Given the MA(2) process $z_t = a_t - 1.2a_{t-1} + .35a_{t-2}$: (a) check whether it is invertible; (b) calculate its autocorrelation structure; (c) write it as an AR(∞) process.

(5) Given the model $z_t = 5 + .9z_{t-1} + a_t + .4a_{t-1}$: (a) calculate its autocorrelation structure; (b) write it in MA(∞) form; (c) write it in AR(∞) form.

(6) Given the process $(1 - B + .21B^2)z_t = a_t - .3a_{t-1}$:(a) check whether it is stationary and invertible: (b) obtain the autocorrelation function; (c) obtain its AR(∞) representation; (d) obtain its MA(∞) representation.

(7) Obtain the autocorrelation function of an ARMA(1,1) process writing it as an MA(∞).

(8) Prove that if we add two MA(1) processes we obtain a new MA(1) process with an MA parameter which is a linear combination of the MA parameters of the two processes, with weighs that are proportional to the quotients between the variances of the innovations of the summands related to the variance of the innovation of the sum process.

Exercises - Time series analysis

INTEGRATED AND LONG-MEMORY PROCESSES

(1) Prove that the sum and the difference of two stationary processes are stationary.

(2) Prove that the model $z_t = a + bt + ct^2 + a_t$, where a_t is a white noise process, becomes a non-invertible stationary process when two differences are taken.

(3) Prove that the autocorrelations of a random walk can be approximated by $\rho(t, t+k) \simeq 1 - \frac{k}{2t}$.

(4) Simulate an ARIMA (0,1,1) process using parameter values $\theta = .4$, .7 and .9, and study the decay of the autocorrelation function of the process.

(5) Simulate the process $\nabla^2 z_t = (1 - .8B)a_t$ and study the decay of the autocorrelation function of the process.

Exercises - Time series analysis

Seasonal ARIMA processes

(1) Find the seasonal coefficients for a quarterly series that follows the model $z_t = 10 + \cos(\pi t/2 + \pi/8) + a_t$.

(2) Prove that the series from above exercise can be modelled using $\nabla_4 z_t = (1 - B^4)a_t$.

(3) We assume that a monthly series follows the model $z_t = 30 + \cos(\pi t/6 + \pi/8) + V_t + a_t$, where a_t is a white noise process with variance σ_a^2 and process V_t verifies $V_t = V_{t-12} + \epsilon_t$, where ϵ_t is a white noise process with variance σ_ϵ^2. Prove that this series follows the ARIMA model $\nabla_{12} z_t = (1 - \Theta B^{12})a_t$, where $\theta \leq 1$.(Suggestion: prove that the process $\epsilon_t + a_t - a_{t-12}$ has an MA(1)$_{12}$ structure and that the autocorrelation of order twelve is - $\sigma_a^2/(\sigma_\epsilon^2 + 2\sigma_a^2)$).

(4) Find the theoretical autocorrelation function of the process $(1 - .4B)w_t = (1 + .5B^{12})$.

(5) Find the theoretical autocorrelation function of the process $(1 - .4B)(1 - .8B^{12})w_t = a_t$.

(6) Find the theoretical autocorrelation function of the process $w_t = (1 - \theta B)(1 - \Theta B^{12})a_t$ and compare it with that of the non-multiplicative process $w_t = (1 - \theta B - \Theta B^{12})a_t$.

Exercises - Time series analysis

FORECASTING WITH ARIMA MODELS

(1) Given the process $z_t = 2 + .8z_{t-1} - .1z_{t-1} + a_t$ and four observations (4, 3, 1, 2.5) generate predictions for 4 periods ahead.

(2) Indicate what will be the long-term forecast generated by the model from the above exercise.

(3) Assuming that the variance of the innovations in above exercise is 2, calculate the confidence intervals for the predictions for one and two steps ahead.

(4) Calculate predictions for $t = 100, 101$ and 102 and the final prediction equation of the MA(2) process $z_t = 5 + a_t - .5a_{t-1}$, knowing that the predictions carried out with information up to $t = 97$ have been: $\widehat{z}_{97}(1) = 5.1$, and $\widehat{z}_{97}(1) = 5.3$, and that we have later observed $z_{98} = 4.9$ and $z_{99} = 5.5$.

(5) Explain the structure of the forecasts generated by the model: $\nabla z_t = 3 + (1 - .7B)a_t$

(6) Explain the long-term structure of the forecasts using the model: $\nabla \nabla_{12} z_t = (1 - .7B)a_t$.

(7) Prove that in the IMA(1,1) process which can be written $z_{T+1} = (1 - \theta) \sum_{j=0}^{\infty} \theta^j z_{T-j} + a_{T+1}$ it is shown that for $k \geq 2$, $\widehat{z}_T(k) = \widehat{z}_T(k-1)$. Notice that $\widehat{z}_T(2) = (1 - \theta)[\widehat{z}_T(1) + \theta z_T + \theta^2 z_{T-1} + ...]$ and replace the expression of $\widehat{z}_T(1)$.

(8) Calculate the predictability of the process $z_t = 2 + .8z_{t-1} - .1z_{t-1} + a_t$.

Exercises - Time series analysis

IDENTIFYING POSSIBLE ARIMA MODELS

(1) An early criterion for determining the number of differences needed to make a series stationary is the Titner Criterion, which consists in differentiating while the variance of the resulting series diminishes and stopping when the variance increases on taking a new difference. Prove that if we start with a stationary series but with first order autocorrelation greater than .5 the variance of the series diminishes when it is differentiated. Suggestion: let x_t denote the original series and $n_t = \nabla x_t$, then $Var(n_t) = 2\sigma_x^2(1 - \rho_1)$.

(2) Identify a model for the airline series.

(3) Identify a model for the series of the Spanish population over 16 years of age.

(4) Justify that a sample of 100 observations generated by the model $(1 - .2B)z_t = a_t$ can be easily identified as one generated by an MA(1) or an ARMA(1,1). Suggestion: express the model as an MA(1) and take into account that the bounds of the coefficients of the ACF and PACF are $T^{-1/2}$.

(5) Check that the error correction model for testing whether the true model is the M1:$(1-.7B)\nabla z_t = a_t$ or the M2: $(1 - .7B)(1 - \phi B)z_t = a_t$, using $|\phi| < 1$, is $\nabla z_t = \alpha z_{t-1} + .7\nabla z_{t-1} + a_t$. Which values of α indicate each of the two models?

(6) Check the equivalence between the condition $\alpha_0 = 1$ and a unit-root in the augmented Dickey-Fuller test obtaining the coefficients α_i as a function of the ϕ, equating powers in both polynomials. Check that $\alpha_p = -\phi_{p+1}$, $\alpha_{p-1} = -\phi_{p+1} - \phi_p$ is obtained, and in general $\alpha_i = -\sum_{j=i+1}^{p+1} \phi_j$; $i \geq 1$ and $\alpha_0 = \phi_1 + ... + \phi_{p+1}$, which confirms that the condition $(1 - \phi_1 - ... - \phi_{p+1}) = 0$, implies the condition $\alpha_0 = 1$.

Exercises - Time series analysis

ESTIMATION AND SELECTION OF ARMA MODELS

(1) Prove that the variance of the estimator of μ for an AR(p) obtained by conditional estimation is greater than the variance of the sample mean of the entire process.

(2) Prove that in an AR(2) process the marginal distribution of the first observation is normal with mean $E[\omega_1] = \mu$ and $Var[\omega_1] = \dfrac{\sigma^2}{1 - \phi_1^2 - \phi_2^2}$ and for the second observation $E[\omega_2|\omega_1] = \mu + \phi_1(\omega_1 - \mu)$ and $Var[\omega_2|\omega_1] = \sigma^2 \dfrac{\left(1 - \phi_1^2\right)}{1 - \phi_1^2 - \phi_2^2}$.

(3) Prove that the ML estimator of $\widehat{\sigma}_a^2$ for an AR(p) process is:

$$\widehat{\sigma}_a^2 = \sum_{t=p+1}^{T} \left(\omega_t - \widehat{\mu} - \Sigma_{i=1}^p \widehat{\phi}_i(\omega_{t-i} - \widehat{\mu})\right)^2 /T$$

(4) Verify that the state space representation of an AR(1) model has $\Omega = \phi$, $H = 1$ and $R = \sigma^2$.

(5) Write the Kalman filter equations to forecast using an AR(1) and verify that they are reduced to $\widehat{z}_{t|t-1} = \phi z_t$ with variance $p_{t|t-1} = \sigma^2$.

(6) Write the state space equations for an MA(1) and prove that $\Omega = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, H=(0,1) and $R = \sigma^2 \begin{bmatrix} 1 & -\theta \\ -\theta & \theta^2 \end{bmatrix}$ are verified.

(7) Using that for square matrices, A and C, and rectangular matrices, B and D, with appropriate dimensions it is verified that $(A+BCD)^{-1} = A^{-1} - A^{-1}B(DA^{-1}B + C^{-1})^{-1}DA^{-1}$, prove that the revision equation of the covariances of the state estimation can be written as $\mathbf{S}_t^{-1} = \mathbf{S}_{\mathbf{t|t-1}}^{-1} + \mathbf{H}_t^{'} \mathbf{V_t}^{-1} \mathbf{H}_t$.

(8) Letting precision denote the inverse variance, justify that the above expression is interpreted as that the final precision is the sum of the initial precision contributed by the last observation.

(9) Write the equations of the Kalman filter for an AR(2) and relate the calculation method of the predictions with that studied in Chapter 8.

Exercises - Time series analysis

Model diagnosis and prediction

(1) Fit a model to the data for unemployment in Spain and analyze the residuals of the model to verify that they are suitable.

(2) Prove that the distribution mean of $\alpha f(\mu_1, \sigma_1^2) + (1 - \alpha)f(\mu_2, \sigma_2^2)$ is $\mu = \alpha\mu_1 + (1 - \alpha)\mu_2$ and the variance $\alpha\sigma_1^2 + (1 - \alpha)\sigma_2^2 + \alpha(\mu_1 - \mu)^2 + (1 - \alpha)(\mu_2 - \mu)^2$.

(3) Use the above result to obtain the mean and variance when we combine predictions from different models.

(4) Prove that if the most probable model is that of less residual variance the intervals constructed by model averaging will be wider than those of a single model.