INSTITUTE OF ACTUARIAL & QUANTITATIVE STUDIES

Subject: Statistical Techniques and Risk Modelling 4

Chapter: Unit 3 & 4

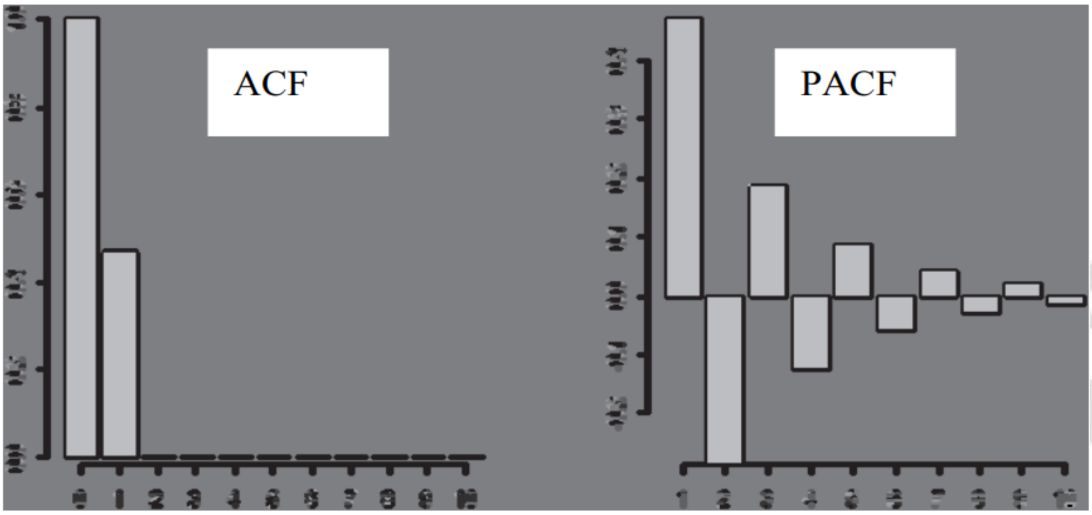
Category: Practice Questions

Note: This contains questions from unit 3 and 4 both, so, some questions may seem to be higher order at the start but after completing 4th unit they will be comfortable.

1. CT6 September 2013 Q9

i) State the three main stages in the Box-Jenkins approach to fitting an ARIMA time series model.

ii) Explain, with reasons, which ARIMA time series would fit the observed data in the charts below.



Now consider the time series model given by

$$X_t = \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \beta_1 e_{t-1} + e_t$$

where e_t is a white noise process with variance σ^2.

iii) Derive the Yule-Walker equations for this model.

iv) Explain whether the partial auto-correlation function for this model can ever give a zero value.

2. CT6 April 2014 Q12

A sequence of 100 observations was made from a time series and the following values of the sample auto-covariance function (SACF) were observed:

Lag	SACF
1	0.68
2	0.55
3	0.30
4	0.06

The sample mean and variance of the same observations are 1.35 and 0.9 respectively.

i) Calculate the first two values of the partial correlation function $\widehat{\phi_1}$ *and* $\widehat{\phi_2}$ [1]

ii) Estimate the parameters (including σ2) of the following models which are to be fitted to the observed data and can be assumed to be stationary.

(a) $Y_t = a_0 + a_1 Y_{t-1} + e_t$

(b) $Y_t = a_0 + a_1 Y_{t-1} + a_2 Y_{t-2} + e_t$

In each case e_t is a white noise process with variance σ^2. [12]

iii) Explain whether the assumption of stationarity is necessary for the estimation for each of the models in part (ii). [2]

iv) Explain whether each of the models in part (ii) satisfies the Markov property. [2]

3. CT6 September 2014 Q9

i) List the main steps in the Box-Jenkins approach to fitting an ARIMA time series to observed data. [3]

Observations $x_1, x_2, \ldots, x_{200}$ are made from a stationary time series and the following summary statistics are calculated:

$$\sum_{i=1}^{200} x_i = 83.7 \qquad \sum_{i=1}^{200} (x_i - \bar{x})^2 = 35.4 \qquad \sum_{i=2}^{200} (x_i - \bar{x})(x_{i-1} - \bar{x}) = 28.4$$

$$\sum_{i=3}^{200} (x_i - \bar{x})(x_{i-2} - \bar{x}) = 17.1$$

ii) Calculate the values of the sample auto-covariances $\hat{\gamma}_0$, $\hat{\gamma}_1$ *and* $\hat{\gamma}_2$. [3]

iii) Calculate the first two values of the partial correlation function $\widehat{\phi_1}$ and $\widehat{\phi_2}$. [3]

The following model is proposed to fit the observed data:

$$X_t - \mu = a_1 (X_{t-1} - \mu) + e_t$$

where et is a white noise process with variance σ^2.

iv) Estimate the parameters μ, a1 and σ^2 in the proposed model. [5]

After fitting the model in part (iv) the 200 observed residual values $\hat{e_t}$ were calculated. The number of turning points in the residual series was 110.

v) Carry out a statistical test at the 95% significance level to test the hypothesis that $\hat{e_t}$ is generated from a white noise process. [4]

4. CT6 April 2015 Q7

The following time series model is being used to model monthly data:

$$Y_t = Y_{t-1} + Y_{t-12} - Y_{t-13} + e_t + \beta_1 e_{t-1} + \beta_{12} e_{t-12} + \beta_1 \beta_{12} e_{t-13}$$

where e_t is a white noise process with variance σ^2.

(i) Perform two differencing transformations and show that the result is a moving average process which you may assume to be stationary. [3]

(ii) Explain why this transformation is called seasonal differencing. [1]

(iii) Derive the auto-correlation function of the model generated in part (i). [8]

5. CT6 S2015 Q11

Consider the following pair of equations:

$$X_t = 0.5X_{t-1} + \beta Y_t + \varepsilon_t^1$$

$$Y_t = 0.5Y_{t-1} + \beta X_t + \varepsilon_t^2$$

where ε_t^1 and ε_t^2 are independent white noise processes.

(i) (a) Show that these equations can be represented as

$$M\begin{pmatrix} X_t \\ Y_t \end{pmatrix} = N\begin{pmatrix} X_{t-1} \\ Y_{t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_t^1 \\ \varepsilon_t^2 \end{pmatrix}$$

where M and N are matrices to be determined.

(b) Determine the values of β for which these equations represent a stationary bivariate time series model. [9]

(ii) Show that the following set of equations represents a VAR(p) (vector auto regressive) process, by specifying the order and the relevant parameters:

$$X_t = \alpha X_{t-1} + \alpha Y_{t-1} + \varepsilon_t^1$$

$$Y_t = \beta X_{t-1} - \beta X_{t-2} + \varepsilon_t^2$$

[3].

6. CT6 April 2016 Q9

Consider the following time series model:

$$Y_t = 1 + 0.6Y_{t-1} + 0.16Y_{t-2} + \varepsilon_t$$

where εt is a white noise process with variance σ^2

(i) Determine whether Y_t is stationary and identify it as an ARMA(p,q) process. [3]

(ii) Calculate E(Y_t). [2]

(iii) Calculate for the first four lags:
- the autocorrelation values $\rho 1$, $\rho 2$, $\rho 3$, $\rho 4$ and
- the partial autocorrelation values $\psi 1$, $\psi 2$, $\psi 3$, $\psi 4$. [7]

[Total 12]

7. CT6 S2016 Q9

In order to model the seasonality of a particular data set an actuary is asked to consider the following model:

where B is the backshift operator and εt is a white noise process with variance σ^2.

The actuary intends to apply a seasonal difference:

$\nabla_s X_t = Y_t$.

(i) Explain why s should be 12 in this case (i.e. $Y_t = X_t - X_{t-12}$). [1]

(ii) Determine the range of values for α and β for which the process will be stationary after applying this seasonal difference. [3]

Assume that after the appropriate seasonal differencing the following sample autocorrelation values for observations of Y_t are $\hat{\rho}_1 = 0$ and $\hat{\rho}_2 = 0.09$.

(iii) Estimate the parameters α and β. [5]

The actuary observes a sequence of observations $x_1, x_2, \ldots, x_T$ of X_t, with $T > 12$.

(iv) Derive the next two forecasted values for next two observations $\hat{x}_{T+1}$ and $\hat{x}_{T+2}$, as a function of the existing observations. [4]

[Total 13]

8. CT6 April 2017 Q6

Model A is a stationary AR(1) model, which follows the equation:

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

where ε_t is a standard white noise process.

i) State two approaches for estimating the parameters in Model A. [2]

Mary, an actuarial student, wishes to revise Model A such that the error terms εt no longer follow a Normal distribution.

ii) Explain which of the approaches in part (i) she should now use for parameter estimation. [2]

iii) Propose a method by which Mary will be able to calculate estimates of the parameters α and σ^2, with reference to any relevant equations. [3]

Mary, has now constructed Model B. She has done this by multiplying both sides of the equation above by (1 − cB), where B is the backshift operator, so that Model B follows the equation:

$$y_t(1-cB) = (\alpha y_{t-1} + \varepsilon_t)(1-cB).$$

iv) Explain why Model A and Model B are identical. [2]

v) Explain for which values of c Model B is stationary. [2]

9. CT6 September 2017 Q10

(i) Show that X_t is not stationary.

Let $\Delta X_t = X_t - X_{t-1}$.

(ii) Show that ΔX_t is stationary.

(iii) Determine the autocovariance values of ΔX_t in terms of those of Y_t.

Now assume that Y_t is an MA(1) process, i.e. $Y_t = \varepsilon_t + \beta\varepsilon_{t-1}$

(iv) Set out an equation for ΔX_t in terms of b, β, ε_t and L, the lag operator.

(v) Show that ΔX_t has a variance larger than that of Y_t.

10. CT6 April 2018 Q7

Consider the following time series model:

$$(1-\alpha B)^3 X_t = \varepsilon_t$$

where B is the backshift operator and et is a standard white noise process with variance σ^2.

i) Determine for which values of α the process Xt is stationary. [2]

Now assume that Xt is stationary.

ii) Calculate the autocorrelation function for the first two lags, ρ1 and ρ2, using the Yule-Walker equations. [7]

iii) State the formulae, in terms of ρ1 and ρ2, for the first two values of the partial auto correlation function φ1 and φ2. [1]

Now assume that α = 1.

iv) Explain how to fit the parameter of this model, given the time series observations X1, X2, …, XT. [2]

11. CT6 September 2018 Q9

An actuary is modelling a set of data which consists of 100 consecutive observations, $y_1, y_2, \dots, y_{100}$. The data has the following statistics:

$$\bar{y} = \frac{1}{100}\sum_{i=1}^{100} y_i = A = 10.5$$

$$\sum_{i=1}^{100} (y_i - \bar{y})^2 = B = 290$$

$$\sum_{i=2}^{100} (y_{i-1} - \bar{y})(y_i - \bar{y}) = C = 60$$

$$\sum_{i=3}^{100} (y_{i-2} - \bar{y})(y_i - \bar{y}) = D = -240$$

(i) Calculate the values of the sample auto-correlations r1 and r2. [3]

(ii) Calculate the first two sample partial auto-correlations values $\hat{\phi}1$ and $\hat{\phi}2$. [2]

The actuary is considering two different models for this data:

Model X: $y_t = a_0 + a_1 y_{t-1} + \varepsilon_t$

Model Y: $y_t = b_0 + b_1 y_{t-1} + b_2 y_{t-2} + \varepsilon_t$

where ε_t is a standard white-noise process, with variance σ^2.

iii) Estimate the parameters (including σ^2) for both Models X and Y, using the method of moments. [10]

(iv) Explain whether each of Models X and Y satisfy the Markov property. [3] [Total 18]

12. CS2 April 2021 Q2

A second order moving average process is defined by the following equation:

$$X_t = \mu + e_t + \beta_1 e_{t-1} + \beta_2 e_{t-2}$$

where e_t is a white noise process with variance σ^2.

(i) Determine $E(X_t)$ and $\mathrm{Var}(X_t)$. [3]

(ii) Determine the autocovariance function, $\{\gamma_k\}$, of X_t for $k \geq 0$. [4]

[Total 7]

13. CS2 April 2021 Q9

Consider the following time series process:

$$Y_t = 1 + 0.3\, Y_{t-1} + 0.1\, Y_{t-2} + e_t$$

where e_t is a white noise process with variance σ^2.

i) Determine whether Yt is stationary and identify the values of p, d and q for which the process is an ARIMA(p,d,q) process. [3]

Let ρk and ϕk denote the values at lag k of the autocorrelation and partial autocorrelation functions, respectively.

ii) Determine the autocorrelation values ρ1, ρ2 and ρ3. [4]

iii) Determine the partial autocorrelation values ϕ1, ϕ2 and ϕ3. [3]

A sample of the process Yt is taken in which the sample autocorrelation values are equal to the theoretical values ρk .

iv) Determine the minimum sample size, n, necessary to reject the null hypothesis of a white noise process, under the Ljung and Box ‘portmanteau’ test using three lags and a 5% significance level. [6]

v) Discuss the relative merits of using a large or a small number of lags in the Ljung and Box ‘portmanteau’ test by considering how the value of n in part (iv) would vary if a different number of lags were used or if the sample autocorrelation values were not equal to the theoretical values. [5]

14. CS2A S2021 Q3

Consider the following time series process:

$$Y_t = e_t + \beta e_{t-1}$$

where e_t is a white noise process with variance σ^2.

(i) Write down the autocorrelation function $\{\rho_k\}$, of Y_t, for $k \geq 0$. [2]

(ii) Determine the possible values of β for which the value of the partial autocorrelation function at lag 2, $\phi_2 = -\frac{1}{3}$. [4]

(iii) Comment on the practical suitability of this time series process for the values of β calculated in part (ii). [2]
[Total 8]

15. CS2 September 2021 Q7

An Actuary is considering using the following process to model a seasonal data set:

$$(1 - B^3)(1 - (\alpha + \beta)B + \alpha\beta B^2)X_t = e_t$$

where B is the backwards shift operator and e_t is a white noise process with variance σ^2.

A seasonal difference series is defined as follows:

$$Y_t = X_t - X_{t-3}$$

(i) Express the equation for the original process Xt in terms of the seasonal difference series, Yt, and the backwards shift operator B. [1]

(ii) Determine the range of values of α and β for which the seasonal difference series, Yt, is stationary. [2]

Let γk and ρk denote the values at lag k of the autocovariance and autocorrelation functions, respectively, of the seasonal difference series, Yt. The first Yule–Walker equation for Yt may be written as follows:

$$1-(\alpha+\beta)\rho_1+\alpha\beta\rho_2=\frac{\sigma^2}{\gamma_0}$$

(iii) Write down the second and third Yule–Walker equations for Yt in terms of ρ1 and ρ2. [2]

The Actuary has observed the following sample autocorrelation values for the series Y_t: $\hat{\rho}_1 = 0.5$ and $\hat{\rho}_2 = 0.2$.

(iv) Estimate, using the equations in part (iii), the parameters α and β based on this information. [5]

[Hint: let M = α + β and N = αβ and use the formula for finding the roots of a quadratic equation.]

(v) Determine the values of the one-step ahead and two-step ahead forecasts, $\hat{x}_{550}$ and $\hat{x}_{551}$, respectively, based on the parameters estimated in part (iv) and the observed values $x_1, x_2, \ldots, x_{549}$ of Xt. [4] [Total 14]

16. CS2A A2022 Q9

A zero-mean, first-order moving average process is defined by the following equation:

$$X_t = e_t + be_{t-1}$$

where et is a sequence of independent and identically distributed N (0, σ^2) random variables.

(i) Derive, in terms of b, the value of p that minimises:

$$E[(X_t - pX_{t-1})^2]$$

[7]

(ii) Comment on your answer to part (i) in the case where b = 1. [3]

(iii) Determine, in the case where b = 1, the values of q and r that minimise:

$$E[(X_t - qX_{t-1} - rX_{t-2})^2].$$

[8]

Total [18]

17. CS2A S2022 Q2

Consider the time series process, Xt, given by:

$$X_t = aX_{t-1} + \frac{1}{2}X_{t-2} + e_t + be_{t-1}$$

where et is a sequence of independent and identically distributed N (0, σ^2) random variables.

Determine the values of the parameters a and b such that Xt is:

(i) stationary. [5]

(ii) invertible. [2]

(iii) I(1). [2]
[Total 9]

18. CS2A A2023 Q5

A sample of size n is taken from a process, Xt, which is believed to be an ARMA(1,1) process of the form

$$X_t = aX_{t-1} + e_t + be_{t-1}$$

where $|a|$, $|b|$ < 1. The sample autocorrelations at lag 1 and lag 2 are 0.65 and 0.325, respectively.

(i) Estimate the parameters a and b by equating the sample autocorrelations to the theoretical values.

Fisher's transformation states that the sample correlation coefficient, r, between two random variables, Y and Z, is such that $\frac{1}{2}\log\left(\frac{1+r}{1-r}\right)$ is approximately Normally distributed with mean $\frac{1}{2}\log\left(\frac{1+\rho}{1-\rho}\right)$ and variance $\frac{1}{n-3}$, where ρ is the theoretical correlation coefficient between Y and Z and n is the sample size.

(ii) Determine the minimum value of n necessary to reject the null hypothesis that b = 0 in favour of the alternative b > 0 at the 95% significance level. You should assume that a is equal to the value determined in part (i) and use Fisher's transformation on the autocorrelation at lag 1. [6] [Total 12]

19. CS2A A2023 Q8

Consider the time-series model:

$$y_t = a\, y_{t-2} + e_t + b\, e_{t-1} \qquad \text{(A)}$$

where et is white noise with mean 0 and variance σ^2.

(i) Derive the possible values of a and b for which the process y_t is stationary and invertible. [4]

(ii) State the values of p and q for which y_t is an ARMA(p, q) process. [1]

If b = 0 the original model (A) reduces to

$$y_t = a\, y_{t-2} + e_t \qquad \text{(B)}$$

(iii) Derive the autocorrelation function for this model while stationarity is assumed to hold. [8]

An actuary attempts to fit the model (A) to some time series data but concludes that the simpler model (B) is more appropriate.

(iv) Discuss how this conclusion could have been reached. [4] [Total 17]

20. CS2A S2023 Q5

Let Xt be the process defined by:

$$X_t = \sum_{i=1}^{t} Y_i$$

where:

$$Y_t = e_t + be_{t-1}$$

where et is a sequence of independent and identically distributed N (0, σ^2) random variables.

(i) State the values of p, d and q for which Xt is an ARIMA(p, d, q) process. [1]

(ii) Demonstrate that $\text{var}(Y_t) = (1 + b^2)\,\sigma^2$ and $\text{cov}(Y_t, Y_{t-1}) = b\sigma^2$. [4]

(iii) Demonstrate that

$$\text{var}(X_t) = t(1 + b^2)\sigma^2 + 2(t - 1)b\sigma^2$$

and that

$$\text{cov}(X_t, X_{t-k}) = (t - k)(1 + b^2)\sigma^2 + (2(t - k) - 1)b\sigma^2$$

for $0 < k < t$. [6]

(iv) Explain what the results in part (iii) imply about the shape of the autocorrelation function of X_t. [2]

[Total 13]