

ETC3550 Exam 2019: solutions

QUESTION 1

Write about a quarter of a page each on any FOUR of the following topics. (Clearly state if you agree or disagree with each statement. No marks will be given without any justification.)

Deduct marks for each major thing missed, and for each wrong statement. In general, be relatively generous if the answer makes sense and contains the main ideas.

(a) *The trouble with statistical methods of forecasting is that they assume the patterns in the past will continue in the future.*

- Statistical methods of forecasting do assume some aspects of the past patterns will continue. However, the better models assume that the way patterns have changed over time will continue to occur into the future. So it is not so much that the past patterns will repeat, but that the past patterns will continue to change at the same rate as they have in the past. 2
- Of course, this may not be true, and then the forecasts will not be good. But for short-term forecasts, it is usually a reasonable assumption. 2
- Some forecasting methods make strong assumptions about the past/future. e.g., linear time trends. Others are highly adaptive and so make relatively weak assumptions. e.g., ETS models. 1

(b) *A time series decomposition into trend, seasonal and remainder terms is only useful when there are no cycles in the data.*

- In time series decomposition, trend and cycle are combined to give the “trend” component. 2
- The decomposition is useful whenever there is some (possibly changing) seasonality in the data. 1
- It can be used for better understanding the data, or for forecasting components separately. 2

(c) *With STL decompositions and ETS models, we always need to transform our data before estimating the components.*

- This is not true. 1
- STL is an additive decomposition and therefore we need to take a transformation before we apply the decomposition if the data has monotonically changing variance. 2
- ETS models can have multiplicative components and therefore we do not need to take a transformation. 2

(d) *Some ETS models are not always suitable and should be avoided.*

- This is true. 1
- We have models that are unstable because we divide through by a state and can cause numerical instabilities, namely $ETS(A,N,M)$, $ETS(A,A,M)$, $ETS(A,A_d,M)$. 2

- Multiplicative error models are only useful for strictly positive data and are numerically unstable for data containing zeros or negative values

2

(e) *The combination of AR and MA components guide long-run ARIMA forecasts.*

- This is not true. The combination of c and d guide these. 1
- If $c = 0$ then for $d = 0$, $d = 1$ and $d = 2$, long-term ARIMA forecasts will respectively go to 0, non-zero constant, follow a straight line. 2
- If $c \neq 0$ then $d = 0$, $d = 1$, $d = 2$, long-term forecasts will respectively go, to the mean of the data, follow straight line, quadratic trend (don't do it). 2

(f) *Linear regression models are simplistic because the real world is nonlinear.*

- True, but that doesn't make linear models not useful. A linear regression model is often a useful but simple approximation to reality that works well. 1
- If there is not enough data to estimate the nonlinearity, particularly if the nonlinear is too mild to estimate well, then a linear model is often the best approach. 2
- The linearity is in the parameters, not in the functional form. So nonlinear relationships can be modelled using a linear model. 2

[Total: 20 marks]

— END OF QUESTION 1 —

QUESTION 2

(a)

- the series is trending. 1
- strong seasonality with seasonal variation increasing with the level of the series, i.e., multiplicative seasonality. 1
- The trough is during summer in the US, July and August. 1
- A bit harder to see but there is another dip in Jan especially visible in the early years. Notice for example Jan 2000 (2001) in the seasonplot is below Dec 1999 (2000). This becomes more visible in the STL decomposition that follows. 1

(b)

- This is an STL decomposition with the three panels showing trend-cycle, seasonal and remainder components. 2
 - as STL is an additive decomposition a log transformation is first applied to account for the multiplicative seasonality. 1
 - although the window is quite large, the seasonal component is changing over time. 1
- It shows that the trough at the beginning of each year due to the Christmas period has disappeared over the years 1

- (c) A. *Seasonal naïve model*. Arguably suitable, the trend may be deemed to flatten towards the end. We'll take both yes and no answers with clear justification. 1
- B. *Drift method plus seasonal dummies*. Not suitable as seasonality is changing. Also better with a log or Box Cox transformation. 1
- C. *Holt-Winters additive damped trend method*. Not suitable, seasonality is multiplicative. 1
- D. *Holt-Winters multiplicative damped trend method*. Perfectly suitable, both multiplicative seasonality and damped trend would be appropriate. 1
- E. *ETS(A,N,M)*. One of the combinations that is not allowed. 1
- F. *ETS(M,A_d,M)*. Suitable for both accounting for changing seasonality and the dampening of the trend. 1
- G. *ARIMA(1,1,4)*. Not suitable - no seas component. 1
- H. *ARIMA(3,1,2)(1,1,0)₁₂*. Possibly suitable as taking seasonal and first order differences. 1
- I. *ARIMA(0,1,1)(2,0,0)₁₂*. Not suitable you would need a seasonal difference. 1
- J. *Regression with time and Fourier terms*. Not suitable, trend not linear (seems to dampen towards the end) and seasonality is changing. 1

[Total: 20 marks]

— END OF QUESTION 2 —

QUESTION 3

(a)

- both models have multiplicative errors, additive trend and multiplicative seasonality 1
- the trend in the second model, damped, is a damped trend 1

(b)

- α for trend is lower than for damped. Hence level for damped is changing more rapidly relative to trend. This is reflected on the estimated level component in Figure 5 being wigglier for damped. 2
- β for trend is 0 hence the slope is not changing at all. This is shown in the slope panel of Figure 5 where the slope is not changing up to the third decimal place — recall that the bar on the right shows the relative scale. So more of a global trend here. 1
- β for damped is quite low (not exactly zero). The combination of this and the dampening parameter is making the slope to slightly change (changes in the second decimal place). 1
- There is a big contrast in the estimated smoothing coefficients for the seasonal components with γ for trend being quite high and hence the seasonal component changing rapidly in Figure 5 while γ for damped is much lower and the seasonal component is changing much slower in the bottom panel of 5. 2

(c)

- both sets of residuals are autocorrelated shown both from the ACF but also the null of white noise is easily rejected by the Ljung-Box test. 1
- histograms look close to normal for both sets. 1
- Point forecasts will be ok but predictions intervals will not. 1
- From the two sets of residuals probably the damped residuals are better given the significant spike at lag 1 for trend. 1

(d)

- Based on the AICc and the overall quality of the residuals I would choose damped. 2

(e) The forecasts from the two models are very different

- trend forecasts include an estimated upwards global trend, although the series seem to flatten out towards the end. 1
- In contrast damped generates flat forecasts due to the effect of the non-zero β and the dampening parameter. 1
- The seasonal components projected from the two models are also very different with trend forecasts including the changed over the estimated sample component reflecting more the seasonality observed towards the end of the sample. 1
- Another stark contrast is the width of the prediction intervals with damped being much wider reflecting more uncertainty. The forecasts, and particularly the wider prediction intervals, make me feel comfortable with the decision to go with the damped ETS(M,Ad,M) model. 1

(f)

$$y_t = (\ell_{t-1} + \phi b_{t-1})s_{t-m}(1 + \varepsilon_t)$$

$$\ell_t = (\ell_{t-1} + \phi b_{t-1})(1 + \alpha \varepsilon_t)$$

$$b_t = \phi b_{t-1} + \beta(\ell_{t-1} + \phi b_{t-1})\varepsilon_t$$

$$s_t = s_{t-m}(1 + \gamma \varepsilon_t)$$

2

[Total: 20 marks]

— END OF QUESTION 3 —

QUESTION 4

(a)

- Row 1: ACF is slow decaying reflecting strong trend and a slightly higher seasonal spike (although not too easy to see). PACF spike at lag 1 close to 1 also showing the strong trend 1
- Row 2: taking first order difference leaves a strong seasonal component reflected by the slow decay in the seasonal spikes in the ACF. Hence stationarity has yet to be achieved. Seasonal differencing is definitely necessary. 1
- Row 3: seasonal differencing seems to get closer to stationarity. The ACF decays at a faster rate than before but still slow and PACF spike is still quite high, although not as close to 1. The time plot of the seasonally differenced series shows that it still jumps around a little. Probably close enough to stationary. A first order difference will ensure stationarity. 1
- Row 4: stationarity seems to have been achieved after both seasonal and first order differencing. 1

(b)

- arima1 relates to the third set of ACF/PACF, where only a seasonal difference has been applied 1
- The AR(1) component captures the dynamics shown in the single spike in the PACF and the decaying spikes in the ACF. 1
- The seasonal components SAR(1) SMA(1) capture the seasonal spikes on both ACF and PACF. 1
- arima2 relates to the fourth set of ACF/PACF, where both a seasonal and a first order difference has been applied 1
- The MA(1) component captures the dynamics shown in the just significant spike at lag 1 in the ACF. 1
- Equivalently the seasonal component SMA(1) captures the seasonal spike on the ACF. 1

(c)

$$(1 - B)(1 - B^{12})y_t = (1 + \theta B)(1 + \Theta B^{12})\varepsilon_t$$

$$(1 - B - B^{12} + B^{13})y_t = (1 + \theta B + \Theta B^{12} + \theta\Theta B^{13})\varepsilon_t$$

$$y_t - y_{t-1} - y_{t-12} + y_{t-13} = \varepsilon_t + \theta\varepsilon_{t-1} + \Theta\varepsilon_{t-12} + \theta\Theta\varepsilon_{t-13}$$

$$y_t = y_{t-1} + y_{t-12} - y_{t-13} + \varepsilon_t + \theta\varepsilon_{t-1} + \Theta\varepsilon_{t-12} + \theta\Theta\varepsilon_{t-13}$$

where $y_t = \log(\text{daycare})$ and ε_t is white noise with mean 0 and variance 0.000122². The estimated parameters are $\theta = -0.142$ and $\Theta = -0.729$. 1

(d)

- Some significant spikes in the ACF due to multiple tests at a 5% level of significance (we expect 1 out of 20 to be incorrectly rejecting the Null, type I error) 1
- The Ljung-Box test clearly shows that the Null white noise ($\rho_1 = \dots = \rho_{24} = 0$) cannot be rejected. 1

(e)

- `arma1` has only one order of differencing and therefore a constant/drift component is required to generate trending forecasts. 1
- `arma2` has two lots of differencing and therefore no constant/drift is allowed nor required to generate trending forecasts. 1
- The drift component in `arma1` projects a global trend/stronger trend in contrast to `arma2` which models 2 lots of differencing. 1
- As expected the two lots of differencing in `arma2` result to wider forecast intervals compared to `arma1`. 1

[Total: 20 marks]

— END OF QUESTION 4 —

QUESTION 5

(a) Because we will then be including more fourier terms than seasons. 2

(b) You would select fit5 as it has min AIC. 1

(c)

$$y_t = a + bt + \sum_{k=1}^5 \left[\alpha_k \sin\left(\frac{2\pi kt}{m}\right) + \beta_k \cos\left(\frac{2\pi kt}{m}\right) \right] + \eta_t$$

$$(1 - \phi B)(1 - \Phi B^{12})\eta_t = (1 + \theta B)(1 + \Theta B^{12})\varepsilon_t$$

where $y_t = \log(\text{daycare})$

and $\varepsilon \sim \text{NID}(0, \sigma^2)$.

- a, b are the intercept and slope coefficients respectively, guiding the long-term linear trend of the forecasts. 1
- $\phi, \theta, \Phi, \Theta$ are the ARIMA coefficients for the error term guiding the short-run dynamics around the long-term linear trend. 1
- α_k, β_k are the coefficients of the Fourier terms guiding the seasonal variation. 1

(d)

$$\exp\left(\log(35.343) \pm 1.96\sqrt{0.000115}\right) = (34.608, 36.094)$$

(e) Although the data is non-stationary, clearly trending, the inclusion of a linear trend has accounted for this returning stationary residuals, which are in turn modelled by a stationary ARMA process. In fact the trend included in the model is exponential given the log transformation on the left hand side (probably will not expect them to comment on this). 2

(f)

- The plot and ACF shows that the residuals are like white noise (with some significant spikes at very long lags). The LB test rejects the null of joint zero autocorrelations for the first 24 lags of the residuals (this is probably due to the spikes at 18 and 21). However, the significant spikes are very small, so will have almost no effect on the point forecasts or prediction intervals. 2
- The histogram shows that the residuals are close to normal, although the tails are a little longer. The prediction intervals may be affected slightly, but probably not much. 1
- The time plot shows that the residuals have mean zero and constant variance. 1

[Total: 20 marks]

— END OF QUESTION 5 —