MS6215: Forecasting Methods for Business; Answers to Exam Questions (2017/18)

1. $98412 \times 2 = 196824$

2. $\frac{36}{2}(196824^2) = 697316727458$

3. TSS=4634264289
   RMSE=9616.10928
   MSE=92469557.68
   SSE=MSE$\times$34=3143964961
   $\therefore R^2 = 1 - \frac{3143964961}{4634264289} = 0.3216$

4. $H_0 : \beta_1 = 0$ vs. $H_1 : \beta_1 \neq 0$
   SSR=1490299328 $\therefore F = \frac{SSR/1}{SSE/34} = \frac{1490299328}{92469557.68} = 16.117$
   $F_{(1,34;0.05)}$=4.10
   $\therefore$ Reject H_0.

5. Based on the data features shown in Figure 1, a positive slope coefficient is expected.
   $t = \sqrt{F} = 4.0146 = \frac{b_1}{154.27796}$
   $\therefore b_1 = 619.364$.

6. $\widehat{Y}_{38} = 86954 + 619.364(38) = 110489.83$

7. Some basic properties of the Z, t, χ^2 and F distributions are in order:

   - Let $Z_i \sim N(0, 1)$ and Z_i's be distributed independently. Then $W = \sum_{i=1}^{v} Z_i^2 \sim \chi^2_{(v)}$
   - Let $Z \sim N(0, 1)$. Provided that Z and W are independent, then $t = Z/\sqrt{W/v} \sim t_{(v)}$.
   - Let $W_1 \sim \chi^2_{(v1)}$ and $W_2 \sim \chi^2_{(v2)}$. Provided that W_1 and W_2 are independent, then $F = \frac{W_1/v1}{W_2/v2} \sim F_{(v1,v2)}$.
   - From the above definitions of t and F, it is straightforward to see that $t^2_{(v)} = F_{(1,v)}$.

   Now, consider the test of $H_0 : \mu = 0$ vs. $H_1 : \mu \neq 0$. The usual statistic for testing H_0 is $T = \frac{\bar{Y}}{s/\sqrt{n}}$, which may be written as $T = \frac{\sqrt{n}\bar{Y}}{\sigma} / \sqrt{\frac{\sum(Y_i - \bar{Y})^2}{\sigma^2(n-1)}}$. Under H_0, the numerator in the second expression of T has the Z distribution and the denominator is the square root of a χ^2 divided by the corresponding d.o.f. (it is well-known that $\bar{Y} \sim N(\mu, \sigma^2/n)$ and $\sum(Y_i - \bar{Y})^2/\sigma^2 \sim \chi^2_{(n-1)}$). As well, the numerator and denominator are independent. Hence, $T \sim t_{(n-1)}$. Now, $t^2_{(n-1)} = F_{(1,n-1)} = \frac{\chi^2_{(1)}/1}{\chi^2_{(n-1)}/(n-1)}$. From the SAS output, $(\frac{n}{2}a_0^2) = 697316727458$. But $\frac{a_0}{2} = \bar{Y}$. Hence $n\bar{Y}^2/\sigma^2 = 348658363724/\sigma^2 \sim \chi^2_1$. Also, $s^2 = \frac{\sum(Y_i - \bar{Y})^2}{n-1} = 4634264289/35$. It is readily seen that $\frac{348658363724}{4634264289/35} = 2633.22 \sim F_{(1,35)}$, and this is precisely the square of the T statistic for testing H_0. Now, $F_{(1,35;0.05)} \approx 4.10$ $\therefore$ Reject H_0.

8. $h = 18$

9. $A = \sqrt{(-8740.25)^2 + (-7835.8)^2}$
   $P = tan^{-1}(\frac{-7835.8}{-8740.25})$
   For the last harmonic wave, $Asin(\pi t + P) = cos(\pi t)Asin(P)$. Hence A and P cannot be separated.

10. $f_3 = 3/36$ $L_3 = 12$

11. $\frac{n}{2}(b_{13}^2 + b_{23}^2) = 18((-8740.25^2) + (-7835.8^2)) = 2480251170$

12. $H_0 : b_{13} = b_{23} = 0$
    $F = \frac{2480251170/2}{(4634264289-2480251170)/33} = 18.99$
    $F_{(2,33;0.05)} \approx 3.3$. $\therefore$ Reject H_0.

    $H_0 : b_{1,12} = b_{2,12} = 0$
    One can similarly work out that
    $F = 3.943$
    Reject H_0.

13. $\widehat{Y}_t = 98412 - 8740.25sin(2\pi\frac{1}{12}t) - 7835.8cos(2\pi\frac{1}{12}t)$

14. $t = 38$ is equivalent to $t = 37$ under Proc Spectra.
    Hence $\widehat{Y}_{37} = 98412 - 8740.25sin(2\pi\frac{1}{12}37) - 7835.8cos(2\pi\frac{1}{12}37) = 87256.0722$

15. Maximum likelihood estimation is applied under Proc Arima. Hence $var(Y_t) = s^2 = 4634264289/36 = 128729563.5$. This leads to

    $r_1 = 0.51976$ $r_2 = 0.30488$ $r_3 = 0.28311$ $r_4 = -0.01415$ $r_5 = -0.26077$ $r_6 = -0.18074$ $r_7 = -0.26868$ $r_8 = -0.16206$ $r_9 = 0.07176$

16. $2std(r_k) = 0.33$. Only ρ_1 is significantly different from zero. The series is stationary as r_1 is not close to 1 and the ACF cuts off after lag 1, i.e., the ACF does not mimic the random walk pattern of $(t-k)/t$.

17. Not a white noise process based on the Q test for autocorrelation check for white noise and the ACF behaviour.

18. $p = q = 0$ $DF = k - p - q = k = 6$.

19. ACF cuts off after 1 (ρ_2 is marginally significant)
    PACF cuts off after lag 1
    These features do not fit in any standard ARMA models (Recall that the ARMA model is only an approximation to the data generating process).
    Could try AR(1), MA(1), ARMA(1,1) as approximations.

20. $H_0 : \delta = 0$ vs. $H_1 : \delta \neq 0$
    $t = \frac{\bar{Y}}{s/\sqrt{n}} = \frac{98412}{11345.91/6} = 52.04 > 2$
    $\therefore$ Reject H_0.

21. The two tests are equivalent and the test statistics differ only in terms of the estimate s^2 substituted for the unknown σ^2. Note that s^2 in Q.20 is the maximum likelihood estimator, whereas s^2 in Q.7 is the unbiased estimator of σ^2. If either the maximum likelihood or the unbiased estimator is used in both questions, then t^2 from Q.20 is equal to F in Q.7.

22. $\widehat{\mu} = \frac{\widehat{\delta}}{1-\widehat{\phi}_1}$. Hence for Table 4, $\widehat{\delta} = 98170(1 - 0.53824) = 45330.98$; for Table 5, $\widehat{\delta} = 98425.6$.

23. For Table 4, $\widehat{Y}_t = 45330.98 + 0.53824Y_{t-1}$
For Table 5, $\widehat{Y}_t = 98425.6 + 0.51818e_{t-1} + 0.09168e_{t-2}$

24. $AIC = -2(ln(L) - g)$; $BIC = -2ln(L) + gln(n)$. Both AIC and BIC are penalised versions of the log-likelihood, which always increases as more lagged terms are added to the model. The penalty term for AIC is $2g$ (i.e., 2 times the number of coefficients), whereas for the BIC, it is $gln(n)$. As $ln(n)$ is usually larger than 2, BIC imposes a larger penalty than the AIC on the log-likelihood when the model's dimension increases. Hence BIC typically favours a more parsimonious model. For Table 4, AIC=766.522, BIC=769.689 For Table 5, AIC=770.26, BIC=775.01

25. Although neither the AR(1) nor the MA(2) results in uncorrelated residuals, the AR(1) model passes the t-test and yields smaller AIC and BIC values; it is superior by all accounts to the MA(2) model.

26. $\rho_1 = \frac{-\theta_1(1-\theta_2)}{1+\theta_1^2+\theta_2^2} = \frac{0.51818(1+0.09168)}{1+0.51818^2+0.09168^2} = 0.444088$
$\rho_2 = \frac{-\theta_2}{1+\theta_1^2+\theta_2^2} = 0.071973$

27. $Y_t = (1 - B)^2 W_t = W_t - 2W_{t-1} + W_{t-2}$
$\therefore W_t - 2W_{t-1} + W_{t-2} = 45330.98 + 0.53824(W_{t-1} - 2W_{t-2} + W_{t-3})$, and hence
$W_t = 45330.98 + 2.53824W_{t-1} - 2.07648W_{t-2} + 0.53824W_{t-3}$

28. ARIMA(1,2,0)

29. $H_0 : \rho = 0$ vs. $H_1 : \rho > 0$
$DW = 1.32$, $k' = 1$, $n = 36$, $d_L = 1.411$, $d_U = 1.525$
As $DW < d_L$, reject H_0.

30. Q.19 and Q.25 suggest that an AR(1) model, which includes Y_{t-1} as a regressor, is an acceptable specification for Y_t. This suggests that autocorrelation in the residuals as detected under Q. 29, is probably due to the omission of Y_{t-1} as an explanatory variable from the model in Table 1.

31. $\alpha = 0.5$. The forecast adjusts reasonably quickly to new observations, but it also depends on the γ value.

32. $\alpha = 1$ and $\gamma = 0 \Rightarrow Y_{t+1} = Y_t + b_0$.

33. As Figure 1 shows, the data series experiences a number of abrupt, sudden changes. This violates one basic assumption of exponential smoothing which requires the data to slowly evolve over time.