Forecasting Homework 1

**1.a.**



Yes, GDP grows almost linearly over time.

**1.b.**



Yes, the log GDP grows linearly over time.

**1.c.**

Regression Analysis: LogGDP versus Time

Analysis of Variance

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Source | DF | Adj SS | Adj MS | F-Value | P-Value |
| Regression | 1 | 119.014 | 119.014 | 26782.21 | 0.000 |
| Time | 1 | 119.014 | 119.014 | 26782.21 | 0.000 |
| Error | 281 | 1.249 | 0.004 |  |  |
| Total | 282 | 120.263 |  |  |  |

Model Summary

|  |  |  |  |
| --- | --- | --- | --- |
| S | R-sq | R-sq(adj) | R-sq(pred) |
| 0.0666617 | 98.96% | 98.96% | 98.94% |

Coefficients

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Term | Coef | SE Coef | T-Value | P-Value | VIF |
| Constant | 7.66121 | 0.00795 | 964.12 | 0.000 |  |
| Time | 0.007938 | 0.000049 | 163.65 | 0.000 | 1.00 |

Regression Equation

|  |  |  |
| --- | --- | --- |
| LogGDP | = | 7.66121 + 0.007938 Time |

Fits and Diagnostics for Unusual Observations

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Obs | LogGDP | Fit | Resid | Std Resid |  |
| 10 | 7.60130 | 7.74059 | -0.13930 | -2.10 | R |
| 11 | 7.61224 | 7.74853 | -0.13629 | -2.06 | R |
| 12 | 7.60323 | 7.75647 | -0.15324 | -2.31 | R |
| 276 | 9.71400 | 9.85211 | -0.13811 | -2.09 | R |
| 277 | 9.71544 | 9.86005 | -0.14461 | -2.18 | R |
| 278 | 9.72098 | 9.86799 | -0.14701 | -2.22 | R |
| 279 | 9.72783 | 9.87593 | -0.14809 | -2.24 | R |
| 280 | 9.73219 | 9.88386 | -0.15167 | -2.29 | R |
| 281 | 9.73526 | 9.89180 | -0.15654 | -2.36 | R |
| 282 | 9.74280 | 9.89974 | -0.15694 | -2.37 | R |
| 283 | 9.75056 | 9.90768 | -0.15712 | -2.37 | R |

*R  Large residual*

Prediction for LogGDP

Regression Equation

|  |  |  |
| --- | --- | --- |
| LogGDP | = | 7.66121 + 0.007938 Time |

Settings

|  |  |
| --- | --- |
| Variable | Setting |
| Time | 285 |

Prediction

|  |  |  |  |
| --- | --- | --- | --- |
| Fit | SE Fit | 95% CI | 95% PI |
| 9.92355 | 0.0079884 | (9.90783, 9.93928) | (9.79140, 10.0557) |

Yes, the prediction interval denoted above is valid since the log GDP grows linearly over time and the values are close to the fit at 95% CI. The PI seems wider than the corresponding CI. This might be because predicting a single response value is less certain than predicting the mean response value.

1.d.

Regression Analysis: LogGDP versus Time

The regression equation is  
LogGDP = 7.661 + 0.007938 Time

Model Summary

|  |  |  |
| --- | --- | --- |
| S | R-sq | R-sq(adj) |
| 0.0666617 | 98.96% | 98.96% |

Analysis of Variance

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Source | DF | SS | MS | F | P |
| Regression | 1 | 119.014 | 119.014 | 26782.21 | 0.000 |
| Error | 281 | 1.249 | 0.004 |  |  |
| Total | 282 | 120.263 |  |  |  |

Fitted Line: LogGDP versus Time



At R-sq of 99% the line seems to fit well.

1.e.



The above plot indicates a positive serial correlation among the error terms. Therefore, assumption of independent error terms is violated. Thus using a linear regression model will spoil the validity of the forecast interval.

2.a.



The above time series plot is not a linear model with high volatility.

2.b.

Actual value for 29-Dec-2017 = 1.2022

# Case 1: Fitting straight line to observations 1 to 700

Prediction for EURO

Regression Equation

|  |  |  |
| --- | --- | --- |
| EURO | = | 1.09976 - 0.000380 Time |

Settings

|  |  |
| --- | --- |
| Variable | Setting |
| Time | 700 |

Prediction

|  |  |  |  |
| --- | --- | --- | --- |
| Fit | SE Fit | 95% CI | 95% PI |
| 0.834059 | 0.0028275 | (0.828508, 0.839610) | (0.760334, 0.907785) |

# Case 2: Fitting straight line to observations 701 to 4772

Prediction for EURO

Regression Equation

|  |  |  |
| --- | --- | --- |
| EURO | = | 1.20613 + 0.000016 Time |

Settings

|  |  |
| --- | --- |
| Variable | Setting |
| Time | 4772 |

Prediction

|  |  |  |  |
| --- | --- | --- | --- |
| Fit | SE Fit | 95% CI | 95% PI |
| 1.28218 | 0.0046179 | (1.27312, 1.29123) | (0.993120, 1.57124) |

We can observe that while for Case 1, the PI is way off than the actual value for 29-Dec-2017 with the range not including the figure. However, in Case 2 the PI does include the actual value with a margin of +/-0.3 on both sides. This can be explained by pointing that predicting a single response value can be less certain than predicting the mean response value.