TOR VERGATA
UNIVERSITÀ DEGLI STUDI DI ROMA

Quantitative Methods III - Mock Exam

Exercise 1

Let $Y = y_i$ and $X = x_i$ be the sample data on cigarette demand (in thousands) and retail price (for 20 cigarettes), observed for five cigarette brands.

X	Y
4.6	15
4.0	20
4.5	27
4.8	29
5.0	31

(a) Estimate the regression parameters.

(b) Draw the scatter plot with the regression line.

(c) Interpret the value of the regression coefficient.

(d) Given that the variance of y is 179.2, estimate the linear determination coefficient R^2.

Exercise 2

Consider the following models, estimated on a sample of 157 monthly observations of the following variables:

- y_t = Microsoft stock excess return
- $ExRm$ = market excess return (S&P500)
- $Inflation$ = inflation growth rate
- M = M3 growth rate
- $Growth$ = GDP growth rate
- $Spread$ = bond risk differential growth rate
- TS = yield spread (10-year vs 3-month rate)
- $Crisis$ = dummy variable (1 from Sep 2008 due to Lehman Brothers collapse until end of 2011)

	(1)	(2)	(3)	(4)	(5)	(6)
Constant	-0.758	-1.12	-0.685	-0.793	-0.826	-0.926
	(0.739)	(0.925)	(0.789)	(0.734)	(0.788)	(0.786)
Ex-Rm	1.19	1.18	1.21	1.20	1.19	1.20
	(0.159)	(0.160)	(0.160)	(0.163)	(0.160)	(0.166)
Crisis		1.01				
		(0.55)				
Inflation			0.448	0.825	1.37	
			(1.76)	(1.76)	(1.81)	
M			0.084	0.008	0.0151	
			(0.029)	(0.029)	(0.031)	
Growth				-1.76		
				(1.11)		
Spread					0.280	0.660
					(5.30)	(5.68)
TS					4.83	5.19
					(2.32)	(2.37)
R^2	0.267	0.277	0.280	0.289	0.268	.....
Adjusted R^2	0.262	0.268	0.261	0.271	0.253	0.267
F-test	56.32	28.267	14.767	15.466	18.656	12.365
RSS	13275	13239	13034	12863	13251	12841

(a) Calculate the 95% confidence interval for the Growth variable coefficient in model (3).

(b) Compute R^2 for model (6).

(c) Test the joint insignificance of Inflation and M in model (6) using the classical F-statistic.

M1

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.08143    0.03485
Y_t_1        0.46832    0.03180
Y_t_2        0.29012    0.03498
Y_t_3        0.09465    0.03495
Y_t_4        0.04444    0.03183
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.009 on 990 degrees of freedom
Multiple R-squared:  0.6987,    Adjusted R-squared:  0.6975
F-statistic: 573.9 on 4 and 990 DF,  p-value: < 2.2e-16
```

M2

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.08481    0.03478
Y_t_1        0.47348    0.03160
Y_t_2        0.30364    0.03362
Y_t_3        0.11548    0.03163
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.009 on 991 degrees of freedom
Multiple R-squared:  0.6981,    Adjusted R-squared:  0.6972
F-statistic: 763.9 on 3 and 991 DF,  p-value: < 2.2e-16
```

Exercise 3

The following figure shows two R outputs for two estimated models (M1 and M2) based on $T = 995$ observations of Y_t.

(a) What models have been estimated?

(b) Perform a hypothesis test for $\beta_4 = 0$ in model M1 at the 5% level, calculating the corresponding p-value. What is the test outcome?

(c) Compute the BIC for model M1, given that the residual sum of squares is 1007.9. Which model would you prefer, knowing that BIC for M2 is 0.042?

(d) Given $(Y_{T-3}, \ldots, Y_T) = (-0.052, -1.483, 1.156, -2.200)$, forecast $Y_{T+1|T}$ using model M1.

(e) Construct a 95% interval using the RMSFE obtained via SER for model M1.

(f) The prediction interval:

- ☐ Always exceeds the confidence interval.
- ☐ Is always smaller than the confidence interval.
- ☐ Coincides with the confidence interval.
- ☐ None of the above.

Exercise 4

Consider the following regression:

$$\Delta\hat{Y}_t = 0.4744 \quad (0.4693) + 0.0006 \quad (0.0006)Y_{t-1}$$

where parentheses contain standard errors.

(a) Compute the Dickey-Fuller (DF) test statistic.

(b) Given critical values of -2.86 and -3.41 at 5% significance for DF with intercept and with intercept + trend, respectively, can we reject non-stationarity at 5%? Which critical value is used and why?

(c) Is the DF test used in its augmented version? Explain.

(d) What would the test outcome be if $\hat{\beta}_1 = -0.006$ instead of 0.0006?

1 Solutions

1. (a) Let the sample data be:

X (Price)	Y (Demand)
4.6	15
4.0	20
4.5	27
4.8	29
5.0	31

We estimate the simple linear regression model:

$$Y = \beta_0 + \beta_1 X + u$$

The least squares estimators are given by:

$$\hat{\beta}_1 = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sum (X_i - \bar{X})^2}$$

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X}$$

Computing:

$$\bar{X} = \frac{4.6 + 4.0 + 4.5 + 4.8 + 5.0}{5} = 4.58$$
$$\bar{Y} = \frac{15 + 20 + 27 + 29 + 31}{5} = 24.4$$

Computing $\hat{\beta}_1$:

$$\sum (X_i - \bar{X})(Y_i - \bar{Y}) = (4.6 - 4.58)(15 - 24.4) + \cdots + (5.0 - 4.58)(31 - 24.4) = 5.94$$
$$\sum (X_i - \bar{X})^2 = (4.6 - 4.58)^2 + \cdots + (5.0 - 4.58)^2 = 0.568$$

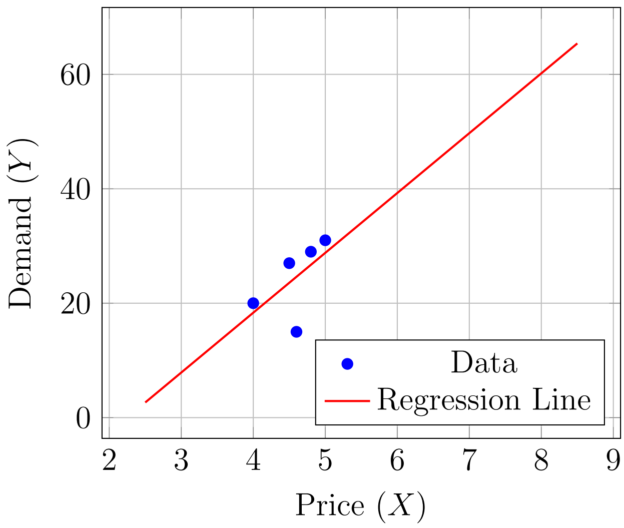
Thus,

$$\hat{\beta}_1 = \frac{5.94}{0.568} = 10.46$$

$$\hat{\beta}_0 = 24.4 - (10.46 \times 4.58) = -23.5$$

So the estimated regression equation is:

$$\hat{Y} = -23.5 + 10.46X$$



(b)

(c) The coefficient $\hat{\beta}_1 = 10.46$ means that for each unit increase in price (X), the demand (Y) increases by approximately 10.46 thousand units.

(d) The coefficient of determination R^2 is given by:

$$R^2 = \frac{ESS}{TSS} = 1 - \frac{RSS}{TSS}$$

Given that the deviance of Y is $TSS = 179.2$

x	y	$\hat{y}$	$u = y - \hat{y}$	u^2
4.6	15	16.5	-1.5	2.25
4.0	20	18.2	1.8	3.24
4.5	27	21.9	5.1	26.01
4.8	29	25.3	3.7	13.69
5.0	31	28.1	2.9	8.41

If we assume $RSS = \sum \hat{u}_i^2 = 117.08$, then:

$$R^2 = 1 - \frac{117.08}{179.2} = 0.3466$$

Thus, 34.6% of the variation in cigarette demand is explained by price.

2. (a) The 95% confidence interval (CI) for a regression coefficient is computed as:

$$\hat{\beta} \pm t_{\alpha/2,n-k} \cdot \text{SE}(\hat{\beta})$$

where:

- $\hat{\beta} = -1.76$ (Growth coefficient in Model 3)
- $\text{SE}(\hat{\beta}) = 1.11$
- $t_{0.025,157-6} = 1.96$ (using the standard ormal distribution)

Computing the interval:

$$-1.76 \pm 1.96 \times 1.11 = -1.76 \pm 2.194 = (-3.954, 0.434)$$

Since the confidence interval contains zero, the Growth variable is not significant at the 5% level.

(b) We aim to calculate the value of R^2 starting from the Adjusted R^2, using the formula:

$$R^2 = 1 - \left(1 - \text{Adjusted } R^2\right) \cdot \frac{N-k}{N-1}$$

Where:

- $N = 157$, the total number of observations.
- $k = 4$, the number of estimated parameters (including the intercept).
- Adjusted $R^2 = 0.267$.

Substituting the values:

$$R^2 = 1 - (1 - 0.267) \cdot \frac{157-4}{157-1} = 0.2813$$

Thus, the R^2 value is 0.2813.

(c) The classical F-statistic for testing joint insignificance of two variables is:

$$F = \frac{(R^2_{unrestricted} - R^2_{restricted})/q}{(1 - R^2_{unrestricted})/(n-k)}$$

- The $R^2_{unrestricted}$ is the value of Model (5) (the same variable of Model (6) plus M and $Inflaction$.
- The $R^2_{restricted}$ is that of Model (6).
- $q = 2$ (number of restrictions)
- $n = 157$, $k = 6$

Computing:

$$F = \frac{(0.289 - 0.2813)/2}{(1 - 0.289)/(157-6)} = 0.817$$

Using an $F_{2,151}$ critical value of approximately 2.35 at 10% significance, we do not reject the null hypothesis. $Inflation$ and M are jointly insignificant in Model (6).

3. (a) AR(4) and AR(3).

(b)

$$t = \frac{0.04444}{0.03183} = 1.396$$

We do not reject the null hypothesis. β_4 is not significant.

(c) We are given the following parameters:

- RSS = 1007.9,
- $T = 995$ (number of observations),
- $k = 5$ (number of parameters estimated: 4 coefficients + 1 intercept).

The BIC is calculated using the formula:

$$\text{BIC} = \ln\left(\frac{\text{RSS}}{T}\right) + \frac{k \cdot \ln(T)}{T}$$

$$\text{BIC} = 0.0129 + 0.0347 = 0.0476$$

Thus, the BIC for the model M_2 is 0.0476.

The BIC of model M_1=0.42, so we prefer M_2 that has a lower BIC.

(d) Given the model $M1$:

$$y_t = 0.08 + 0.46y_{t-1} + 0.29y_{t-2} + 0.09y_{t-3} + 0.04y_{t-4}$$

We are given the values $(Y_{T-3}, Y_{T-2}, Y_{T-1}, Y_T) = (-0.052, -1.483, 1.156, -2.200)$.

To forecast $Y_{T+1|T}$, we substitute these values into the model:

$$Y_{T+1|T} = 0.08 + 0.46 \cdot Y_T + 0.29 \cdot Y_{T-1} + 0.09 \cdot Y_{T-2} + 0.04 \cdot Y_{T-3}$$

Substituting the given values:

$$Y_{T+1|T} = 0.08 + 0.46 \cdot (-2.200) + 0.29 \cdot 1.156 + 0.09 \cdot (-1.483) + 0.04 \cdot (-0.052)$$

Thus, the forecast for $Y_{T+1|T}$ is -0.733.

(e)

$$95\% CI(Y_{T+1|T}) = [-0.733 \pm 1.96 \cdot 1.009] = [-2.709, 1.243]$$

4. Consider the following regression:

$$\hat{\Delta Y_t} = 0.4744 \quad (0.4693) + 0.0006 \quad (0.0006) Y_{t-1}$$

where parentheses contain standard errors.

(a) The Dickey-Fuller test statistic is computed as the ratio of the estimated coefficient on Y_{t-1} to its standard error. Specifically, the DF test statistic is:

$$\text{DF} = \frac{\hat{\beta}_1}{\text{SE}(\hat{\beta}_1)} = \frac{0.0006}{0.0006} = 1$$

Therefore, the Dickey-Fuller test statistic is:

$$\text{DF} = 1$$

(b) The DF test with an intercept does not include a trend, so we should use the critical value for the DF test with an intercept, which is -2.86 at the 5% significance level.

Since the test statistic 1 is greater than the critical value of -2.86, we cannot reject the null hypothesis of non-stationarity at the 5% significance level.

$$\text{DF statistic} = 1 > -2.86 \quad \text{(fail to reject the null hypothesis)}$$

Thus, we cannot reject non-stationarity at the 5% level.

(c) The augmented version of the Dickey-Fuller test includes additional lags of ΔY_t in the regression model to account for autocorrelation in the error terms. In the given regression, only one lag of Y_{t-1} is included, so the test is not augmented.

Since there is no inclusion of lagged differences of ΔY_t, this is the standard (not augmented) Dickey-Fuller test.

(d) If $\hat{\beta}_1 = -0.006$, the test statistic would change. The DF test statistic would be:

$$\text{DF} = \frac{-0.006}{0.0006} = -10$$

With this new test statistic of -10, which is much smaller than the critical value of -2.86, we would reject the null hypothesis of non-stationarity at the 5% significance level.

$$\text{DF statistic} = -10 < -2.86 \quad \text{(reject the null hypothesis)}$$

Thus, the outcome of the test would be different, and we would reject non-stationarity at the 5% level.