# Assignments: Chapter 3(the third lesson)

**3.8** The data file retail lists total U.K. (United Kingdom) retail sales (in billions of pounds) from January 1986 through March 2007. The data are not "seasonally adjusted," and year 2000 = 100 is the base year.

**(a)** Display and interpret the time series plot for these data. Be sure to use plotting symbols that permit you to look for seasonality.

**(b)** Use least squares to fit a seasonal-means plus linear time trend to this time series. Interpret the regression output and save the standardized residuals from the fit for further analysis.

**(c)** Construct and interpret the time series plot of the standardized residuals from part (b). Be sure to use proper plotting symbols to check on seasonality.

**3.14** (Continuation of Exercise 3.8) The data file retail contains U.K. monthly retail sales figures.

**(a)** Obtain the least squares residuals from a seasonal-means plus linear time trend model.

**(b)** Perform a runs test on the standardized residuals and interpret the results.

**(c)** Calculate and interpret the sample autocorrelations for the standardized residuals.

**(d)** Investigate the normality of the standardized residuals (error terms). Consider histograms and normal probability plots. Interpret the plots.

<u>Additional problem.</u> Data $\{y_t\}_{t=1}^{T}$ are drawn from the stationary time series

$$y_t = \beta_0 + \beta_1 x_t + u_t,$$

where $u_t's$ are weak white noises, i.e., $\mathrm{cov}(u_t, x_t) = 0$. Ordinary Least square (OLS) regression method can be used to obtain the estimate of $\beta_0$, $\beta_1$. So, we minimize

$$\sum_{t=1}^{T}(y_t - \beta_0 - \beta_t x_t)^2.$$

Let the resulting estimation of $\beta_0$, $\beta_1$ be $\hat{\beta}_0$, $\hat{\beta}_1$. Then $\hat{y}_t = \hat{\beta}_0 + \hat{\beta}_1 x_t$ is the fitted value of $y_t$, $\hat{u}_t = y_t - \hat{y}_t = y_t - \hat{\beta}_0 - \hat{\beta}_1 x_t$ is the residuals. Denote by $\bar{\mathrm{x}} = \frac{1}{T}\sum_{t=1}^{T}\mathrm{x}_t$, $\bar{\mathrm{y}} = \frac{1}{T}\sum_{t=1}^{T}\mathrm{y}_t$, $\bar{\hat{y}} = \frac{1}{T}\sum_{t=1}^{T}\hat{y}_t = \frac{1}{T}\sum_{t=1}^{T}(\hat{\beta}_0 + \hat{\beta}_1 x_t)$.

**Questions:**

(1) . To prove the algebraic properties of the above OLS regression satisfies

$$\sum_{t=1}^{T} \hat{u}_t = 0, \quad \sum_{t=1}^{T} x_t \hat{u}_t = 0, \ \bar{y} = \bar{\hat{y}}$$

(Hints). To prove the above conclusions from the first order of derivatives of $\sum_{t=1}^{T}(y_t - \beta_0 - \beta_t x_t)^2$

(2) . It is known that $\mathrm{SST} = \frac{1}{T}\sum_{t=1}^{T}(y_t - \bar{y})^2$, $\mathrm{SSR} = \frac{1}{T}\sum_{t=1}^{T}(\hat{y}_t - \bar{\hat{y}})^2$, $\mathrm{SSE} = \frac{1}{T}\sum_{t=1}^{T}(y_t - \hat{y}_t)^2$ are the total sum of squares ( proportional to the variance of the data), the explained sum of squares, the residual sum of squares, respectively. The R-square(Goodness of Fit) is defined as

$$R^2 = \frac{SSR}{SST} = 1 - \frac{SSE}{SST} \qquad (1)$$

Another definition of R-square is

$$R^2 = \frac{\sum_{t=1}^{T}(y_t - \bar{y})(\hat{y}_t - \bar{\hat{y}})}{(\sum_{t=1}^{T}(y_t - \bar{y})^2)\sum_{t=1}^{T}(\hat{y}_t - \bar{\hat{y}})^2}, \qquad (2)$$

which is asymptotically equal to the squared correlation coefficient between the actual and the fitted value of dependent variable, i.e

$$\rho^2(y_t, \hat{y}_t) = \frac{\mathrm{cov}^2(y_t, \hat{y}_t)}{var(y_t)var(\hat{y}_t)}$$

Please prove that the two definitions of (1)-(2) are asymptotically equal (the conclusions of (1) may be useful).