Econ 301 | **Bilkent University**
Econometrics | **Department of Economics**

<u>Sample questions</u>

Work on the exercises 8.1, 7.2, 7.7, 6.5, 4.10 in the text book.

1. The relationship between Unemployment rate and Inflation rate is examined using quarterly data for the period, 1962:1 – 1967:4; (24 observations). The parameters of the relationship are thought to vary between quarters and the researcher defined to following model:

$$Unemp_t = \beta_1 + \delta_1 S_{1t} + \delta_2 S_{2t} + \delta_3 S_{3t} + \beta_2 INF_t + \varepsilon_t$$

where $Unemp_t$ and INF_t are the Unemployment rate and Inflation rate respectively. The seasonal dummies are defined as S_{1t} takes the value of 1 for the first-quarter observations and 0 otherwise; S_{2t} takes the value of 1 for the second-quarter observations and 0 otherwise; S_{3t} takes the value of 1 for the third-quarter observations and 0 otherwise. The estimation results are given in the following table:

```
============================================================
LS // Dependent Variable is UNEMP
Date: 12/06/04   Time: 22:52
Sample: 1962:1 1967:4
Included observations: 24
============================================================
      Variable      Coefficient Std.Error t-Statistic  Prob.
============================================================
         C            7.874718   0.112034   70.28845   0.0000
         S1          -0.029751   0.069246  -0.429646   0.6723
         S2          -0.134887   0.069452  -1.942165   0.0671
         S3          -0.163719   0.069295  -2.362642   0.0290
        INF          -0.211747   0.006710  -31.55472   0.0000
============================================================
R-squared             0.981417    Mean dependent var 4.783333
Adjusted R-squared    0.977504    S.D. dependent var 0.796086
S.E. of regression    0.119402    Akaike info criter-4.067475
Sum squared resid     0.270878    Schwarz criterion -3.822047
Log likelihood        19.75518    F-statistic        250.8546
Durbin-Watson stat    1.407554    Prob(F-statistic)  0.000000
```

a. According to economic theory what is the expected sign for the coefficient of Inflation and why?
b. Write the predicted (fitted) equation for Unemployment ($\hat{Unemp}_t$) in the fourth quarter according to the above estimation results?
c. Is inflation rate variable a significant variable in explaining unemployment? Test the hypothesis.
d. Test the hypothesis that there is no statistical difference between the first quarter and fourth quarter unemployment equations.
e. Test the hypothesis that there is no statistical difference between the second quarter and third quarter unemployment equations. (COVARIANCE matrix is necessary)
f. Test the hypothesis that seasons affect the relationship between unemployment rate and inflation.

[For each hypothesis testing, state the null and alternative, the test statistic, the critical value, your conclusion and your interpretation]

```
============================================================
LS // Dependent Variable is UNEMP
Date: 12/06/04   Time: 23:08
Sample: 1962:1 1967:4
Included observations: 24
============================================================
      Variable      CoefficienStd. Errort-Statistic  Prob.
============================================================
         C           7.765891   0.107917   71.96140   0.0000
        INF         -0.209866   0.007353  -28.54128   0.0000
============================================================
R-squared            0.973703    Mean dependent var 4.783333
Adjusted R-squared   0.972508    S.D. dependent var 0.796086
S.E. of regression   0.131997    Akaike info criter-3.970297
Sum squared resid    0.383311    Schwarz criterion -3.872126
Log likelihood       15.58904    F-statistic        814.6044
Durbin-Watson stat   1.586513    Prob(F-statistic)  0.000000
============================================================
```

```
============================================================
LS // Dependent Variable is UNEMP
Date: 12/06/04   Time: 23:09
Sample: 1962:1 1967:4
Included observations: 24
============================================================
      Variable      CoefficienStd. Errort-Statistic  Prob.
============================================================
         C           4.783333   0.162500   29.43581   0.0000
============================================================
R-squared            0.000000    Mean dependent var 4.783333
Adjusted R-squared   0.000000    S.D. dependent var 0.796086
S.E. of regression   0.796086    Akaike info criter-0.415321
Sum squared resid    14.57633    Schwarz criterion -0.366236
Log likelihood      -28.07067    Durbin-Watson stat 0.047920
============================================================
```

2 . Suppose for explaining the response of variable Y_i we have a set of four explanatory variables, namely X_2, X_3, X_4 and X_5. The following two models are considered estimated with 20 observations:

Model 1: $Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i} + \beta_5 X_{5i} + \varepsilon_i$;

Model 2: $Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \varepsilon_i$:

When the models are estimated the following results are obtained:

$$\widehat{Y}_i = 14 - 0.642 X_{2i} + 0.396 X_{3i} \quad ; R^2 = 0.837, \quad \widehat{s} = 3.072$$

$$\widehat{Y}_i = 14.6 - 0.611 X_{2i} + 0.439 X_{3i} - 0.08 X_{4i} - 0.064 X_{5i} ; R^2 = 0.845, \quad \widehat{s} = 3.190$$

a) Discuss briefly why we can not use R^2 to compare the two models.
b) Compute the SSR, (Sum of Squared Residuals) for both models and SST (sum of squared total).
c) Compute $\overline{R}^2$, (adjusted R^2) for both models. Use this to decide which model is better.

d) Use of formal test, ie a F –test to check whether the variables X_4 and X_5 contribute sufficiently in model 1.

3 .Short-answers

Choose **three** questions from the following list.

a. Define and illustrate graphically a distribution that is **consistent** and a distribution that is **inconsistent**

b. For sample with each observation described by the following equation $Y_t = \beta_1 + \beta_2 X_t + \varepsilon_t$ and a population regression equation, $E(Y_t) = \beta_1 + \beta_2 X_t$, and the sample regression equation, $Y_t = \widehat{\beta}_1 + \widehat{\beta}_2 X_t$, **describe** and **show graphically**, the concepts of Total Sum of Squares, (SST), Explained Sum of Squares (SSE) and Residual Sum of Squares (SSR) and R^2.

c. In the regression $\log(Y_i) = \alpha + \beta \log(X_i) + \varepsilon_i$, how will you define the elasticity of Y with respect to X.

d. Given the following regression equations which can be compared using R^2 statistics to choose the best regression and why? Explain.

   i. $Y_t = \beta_1 + \beta_2 X_t + \varepsilon_t$ estimated with 50 observations

   ii. $Y_t = \beta_1 + \beta_2 X_t + \gamma Z_i + \varepsilon_t$ estimated with 25 observations.

   iii. $\log(Y_t) = \beta_1 + \beta_2 \log(X_t) + \varepsilon_t$ estimated with 50 observations

   iv. $Y_t = \beta_1 + \beta_2 X_t + \gamma(1/Z_i) + \varepsilon_t$ estimated with 25 observations.

e. When can you use adjusted R^2 instead?

4. Consider the model $S_t = \alpha + \beta Y_t + \delta A_t + \lambda P_t + e_t$ where S_t is the sales of a firm in district t, Y_t is the total income in the state, and A_t is the amount of money spent by the company advertising and P_t is the population in that district $(t = 1,2,...,50)$. You suspect that the random error term e_t is heteroscedastic with variance σ^2 that is related to the Advertising expenditure A_t.

Assume that you have the following information about the error structure. **For each of the following cases**, explain how you would revise the estimation technique to obtain estimates that are BLUE. Write an hypothetical dependent variable and independent variable set for each assumption and **prove** that errors of the transformed model is acually homoskedastic.

i) $\text{var}(e_t) = \sigma_t^2 = \sigma^2 A_t$.

ii) $\text{var}(e_t) = \sigma_t^2 = \sigma^2 (1/A_t^4)$

iii) $\text{var}(e_t) = \sigma_t^2 = \sigma^2 (P_t^2 + A_t^2)$

iv) $std(e_t) = \sigma_t = \sigma A_t$

5. Consider the model which explains logarithm of output, LnY_t as a function of logarithm of labor, LnL_t and logarithm of capital, LnK_t inputs:

$$LnY_t = \beta_1 + \beta_2 LnL_t + \beta_3 LnK_t + \varepsilon_t .$$

Suppose that least square estimation on **25** observations on these variables yield the following results:

$$L\hat{n}Y_t = 0.415 + 1.194 LnL_t + 0.217 LnK_t$$

$$s^2 = 0.0076818 \qquad\qquad R^2 = 0.9451$$

$$V\hat{a}r(\hat{\beta}) = \begin{bmatrix} 0.0232 & -0.0354 & 0.0124 \\ -0.0354 & 0.0692 & -0.0294 \\ 0.0124 & -0.0294 & 0.0140 \end{bmatrix}$$

a. Find the 95% interval estimate for β_2.

b. Use a t-test to test the hypothesis $H_0: \beta_2 \geq 1$ against the alternative hypothesis $H_1: \beta_2 < 1$.

c. Test the hypothesis that $H_0: \beta_2 + \beta_3 = 1$ against the alternative of not equal to one.

d. Find the total variation, unexplained variation and explained variation for this model.

e. Test the hypothesis that $H_0: \beta_2 = \beta_3 = 0$ against the alternative of not equal to zero

[Hint: SSResidual of the model $LnY_t = \beta_1 + \varepsilon_t$ **is equal to the** SST of $LnY_t = \beta_1 + \beta_2 LnL_t + \beta_3 LnK_t + \varepsilon_t$ **].**