Assignment No. 2

QUANTITATIVE TECHNIQUES

Question 1

The office of the Registrar took a random sample of 427 students and obtained their grade point averages in college (COLGPA), high school (HSGPA), verbal scores in the SAT (VSAT) and math scores in SAT (MSAT). The following model was estimated (subscript *t* is omitted for simplicity).

$$\textbf{COLGPA} = \beta_1 + \beta_2\,\textbf{HSGPA} + \beta_3\,\textbf{VSAT} + \beta_4\,\textbf{MSAT} + \textbf{u}$$

a. The unadjusted R^2 was 0.22. Because this is very low, we might suspect that the model is inadequate. Test the model for the overall goodness of fit (using a 1 percent level of significance). Be sure to state the null and alternative hypotheses, the test statistic and its distribution, and the criterion for acceptance or rejection. What is your conclusion?

ANSWER:

First we test the hypothesis:

H_0: $\beta_2 = \beta_3 = \beta_4 = 0$
H_1: H_0 is not true and not all slope coefficients are zero.

Next we test the level of significance where $\alpha = 0.01$ which is a given value.

Degree of Freedom numerator = k-1
= 4-1
=3

Degree of Freedom denominator = n-k
= 437-4
= 423

With these values we then test the statistics using the equation as follows.

$$F = \frac{\frac{R^2}{k-1}}{\frac{(1-R^2)}{(n-k)}}$$

It should be noted that a specific decision rule will be used. We reject H_0 if the calculated F value is greater than the tabulated F value. Therefore, if calculated F is greater than 1.7816 we will reject H_0.

Calculating using above equation:

$$F = \frac{\frac{R^2}{k-1}}{\frac{(1-R^2)}{(n-k)}}$$

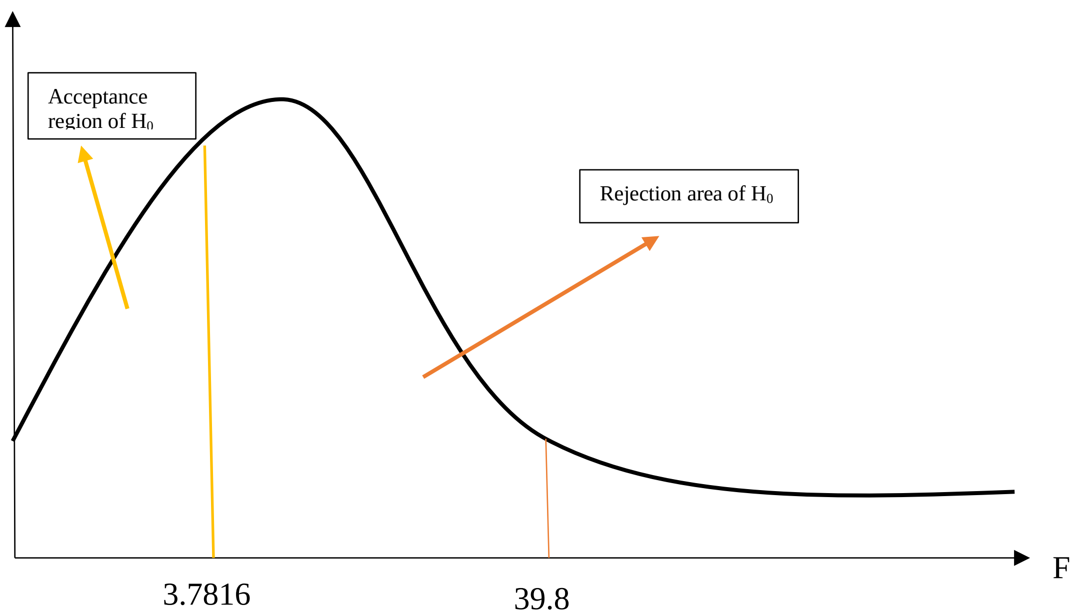
$$F = \frac{\frac{0.22}{3}}{\frac{(1-0.22)}{(423)}}$$

$$= \frac{0.073}{\frac{.78}{423}}$$

$$= \frac{0.073}{0.00184}$$

$$F\ calculated = 39.8$$

Based on the results obtained, we will reject the H0 because our F calculated value is greater than the F tabulated value (39.8 > 3.7816). Based on the results, the null hypothesis is rejected and it is concluded that at minimum one of the regression coefficients is non zero.



b. Test each regression coefficient for significance at the 1 percent level against the alternative that the coefficient is positive. Are any of them insignificant? In the table fill in the test statistic and indicate whether the coefficient is significant or insignificant. State their distribution and degrees of freedom.

ANSWER:

For the particular questions we are testing individual regression coefficients. Hence we will use the T test for testing the hypothesis about individual partial regression co-efficient.

H0: $\beta_1 \le 0$, H1: $\beta_1 > 0$
H0: $\beta_2 \le 0$, H1: $\beta_2 > 0$
H0: $\beta_3 \le 0$, H1: $\beta_3 > 0$
H0: $\beta_4 \le 0$, H1: $\beta_4 > 0$

Using these we will test for significance with $\alpha = 0.01$, one tailed t test.

The following formulas are used;

$$\hat{\beta}_1 = t = \frac{\hat{\beta}_1 - \beta_1}{Se\ (\hat{\beta}_1)}$$

$$\hat{\beta}_2 = t = \frac{\hat{\beta}_2 - \beta_2}{Se\ (\hat{\beta}_2)}$$

$$\hat{\beta}_3 = t = \frac{\hat{\beta}_3 - \beta_3}{Se\ (\hat{\beta}_3)}$$

$$\hat{\beta}_4 = t = \frac{\hat{\beta}_4 - \beta_4}{Se\ (\hat{\beta}_4)}$$

The decision rule used for this is as follows, reject the null hypothesis (H0) if the calculated t is greater than the tabulated t or t-critical. In order to calculate this the t-critical value at 0.01 is 2.33. The test statistics used for this calculation is t-distribution test.

For $\boldsymbol{\widehat{\beta}_1}$;

$$\hat{\beta}_1 = t = \frac{\hat{\beta}_1 - \beta_1}{Se\ (\hat{\beta}_1)}$$

$$t = \frac{0.423 - 0}{0.220}$$

$$\boldsymbol{t = 1.92}$$

For $\widehat{\boldsymbol{\beta}}_{\mathbf{2}}$;

$$\hat{\beta}_2 = t = \frac{\hat{\beta}_2 - \beta_2}{Se\ (\hat{\beta}_2)}$$

$$t = \frac{0.398}{0.061}$$

$$\boldsymbol{t = 6.52}$$

For $\widehat{\boldsymbol{\beta}}_{\mathbf{3}}$

$$\hat{\beta}_3 = t = \frac{\hat{\beta}_3 - \beta_3}{Se\ (\hat{\beta}_3)}$$

$$t = \frac{7.375\ \times 10^{-4}}{2.807\ \times\ 10^{-4}}$$

$$\boldsymbol{t = 2.63}$$

For $\widehat{\boldsymbol{\beta}}_{\mathbf{4}}$

$$\hat{\beta}_4 = t = \frac{\hat{\beta}_4 - \beta_4}{Se\ (\hat{\beta}_4)}$$

$$t = \frac{0.001015}{2.936\ \times\ 10^{-4}}$$

$$\boldsymbol{t = 3.46}$$

Completed Table:

Coefficient	Standard error	Test Statistic	Significance
$\boldsymbol{\beta_1}$ **0.423**	0.220	t-test/distribution	**Insignificant**
$\boldsymbol{\beta_2}$ **0.398**	0.061	t-test/distribution	**Significant**
$\boldsymbol{\beta_3}$ **7.375e-04**	2.807e-04	t-test/distribution	**Significant**
$\boldsymbol{\beta_4}$ **0.001015**	2.936e-04	t-test/distribution	**Significant**

Based on the results as presented in the table above, values of β_2, β_3, β_4 are greater than 2.33 (critical value). This is not true for the constant β_1 which is 1.92, below the critical value. It is concluded that the coefficients of β_2, β_3, β_4 are all significant at the at the one percent level (α=0.01), except for the constant of β_1.

c. Based on your results, write down the names of variables that are candidates for omission from the model.

d. Suppose a student took a special course to improve her SAT scores and increased the verbal and math scores by 100 points each. On average, how much of an increase in college GPA could she expect?

ANSWER:

$$\Delta\ \mathrm{COLGPA} = 0.0007375\ \Delta\mathrm{VSAT} + 0.001015\ \Delta\mathrm{MSAT}$$
$$= 0.07375 + 0.1015$$
$$= 0.17525.$$

Thus, the expected average increase in COLGPA is 0.175.

e. Suppose you want to test the hypothesis that the regression coefficients for VSAT and MSAT are equal. How will you test this hypothesis? [Suppose that cov (β_3, β_4) = -0.25].

ANSWER:

The general unrestricted model is

$$\mathbf{COLGPA = \beta_1 + \beta_2 HSGPA + \beta_3 VSAT + \beta_4 MSAT + u}$$

The marginal effect of VSAT is β_3 and the marginal effect of MSAT is β_4. The test is therefore $\beta_3 = \beta_4$. The alternative is that these two coefficients are unequal. Hence our hypotheses for the test are as follows,

$$\mathbf{H_0: \beta_3 = \beta_4 \text{ or } \beta_3 - \beta_4 = 0}$$
$$\mathbf{H_1: \beta_3 \neq \beta_4 \text{ or } \beta_3 - \beta_4 \neq 0}$$

The variance of the estimated differences of $\hat{\beta}_3 - \beta_4$ is given by the following;

$$Var\left(\hat{\beta}_3\right) + Var\left(\hat{\beta}_4\right) - 2\ Cov\ (\hat{\beta}_3, \hat{\beta}_4)$$

Using the direct two tailed t-test method the calculated t-statistic is therefore as follows,

$$t_c = \frac{\hat{\beta}_3 - \hat{\beta}_4}{\sqrt{Var\left(\hat{\beta}_3\right) + Var\left(\hat{\beta}_4\right) - 2Cov\left(\hat{\beta}_3, \hat{\beta}_4\right)}}$$

For this test, if the numerical value of t_c is greater than t^*_{n-k} (significance level of 2), then our H_0 hypothesis is rejected.

The following are the calculations;

$$t_c = \frac{(\hat{\beta}_3 - \hat{\beta}_4) - (\beta_3 - \beta_4)}{Se(\hat{\beta}_3 - \hat{\beta}_4)}$$

Equation 0-1

$$t_c = \frac{\hat{\beta}_3 - \hat{\beta}_4}{\sqrt{Var\left(\hat{\beta}_3\right) + Var\left(\hat{\beta}_4\right) - 2Cov\left(\hat{\beta}_3, \hat{\beta}_4\right)}}$$

Equation 0-2

$$Se(\hat{\beta}_3 - \hat{\beta}_4) = \sqrt{Var\left(\hat{\beta}_3\right) + Var\left(\hat{\beta}_4\right) - 2Cov\left(\hat{\beta}_3, \hat{\beta}_4\right)}$$

$$Se(\hat{\beta}_3 - \hat{\beta}_4) = \sqrt{7.879 \times 10^{-8} + 8.62 \times 10^{-5} - 2(0.25)}$$

$$= \sqrt{.50}$$

$$Se(\hat{\beta}_3 - \hat{\beta}_4) = 0.707$$

Now we place the value obtained from equation 2 into equation 1.

$$t_c = \frac{(\hat{\beta}_3 - \hat{\beta}_4) - (\beta_3 - \beta_4)}{Se(\hat{\beta}_3 - \hat{\beta}_4)}$$

$$= \frac{(0.0007375 - 0.001015) - (0)}{0.707}$$

$$t_c = -0.0003925$$

Based on our results, the t-tabulated value is greater than the t-calculated value. Therefore, we must accept the H_0 hypothesis at the level of one percent significance concluding that H_0: $\beta_3 = \beta_4$ or β_3 - β_4 =0.

Question 2

The following is known as the **transcendental production function** (TPF), a generalization of the well-known Cobb–Douglas production function:

$$Y_i = \beta_1 L^{\beta_2} K^{\beta_3} e^{\beta_4 L + \beta_5 K}$$

where *Y* = output, *L* = labor input, and *K* = capital input.

After taking logarithms and adding the stochastic disturbance term, we obtain the stochastic TPF as

$$\ln Y_i = \beta_0 + \beta_2 \ln L_i + \beta_3 \ln K_i + \beta_4 L_i + \beta_5 K_i + u_i$$

where $\beta_0 = \ln \beta_1$.

a. For the TPF to reduce to the Cobb–Douglas production function, what must be the values of $\beta 4$ and $\beta 5$?

ANSWER:

The function allows the marginal products of labour and capital to increase before they experience a downfall that is bound to happen. In case of a standard Cobb-Douglas production function, the marginal products decrease from the beginning. Hence, this function allows for the experience of variable elasticity of substitution, then compared to the commonly used Cobb-Douglas model. Hence, the values of β_4 and β_5 must be zero for the TPF to reduce the Cobb-Douglass production function, in this case the value of e^0 is equal to 1 making it the standard model.

b. If you had the data, how would you go about finding out whether the TPF reduces to the Cobb–Douglas production function? What testing procedure would you use?

ANSWER:

If data were available, the testing procedure that will he used to see if TPF reduces to the Cobb-Douglas production function is through the use of the F test f restricted least squares. The following is a general outline of the procedure that would be used with the availability of data.

When conducting the multiple regression F test the following hypotheses will be used in which.

$H_0 = \beta_4 = 0, \beta_5 = 0$
H_1= rejection of H_0 as it is untrue

After establishing this an un-restricted regression is run on

$$\ln Y_i = \beta_0 + \beta_2 \ln L_i + \beta_3 \ln K_i + \beta_4 L_i + \beta_5 K_i$$

Afterwards, RSS_{UR} (unrestricted) is calculated before proceeding to the calculations of restricted regression by including β_4 and β_5 into the equation as follows,

$$\ln Y_i = \beta_0 + \beta_2 \ln L_i + \beta_3 \ln K_i$$

Then its RSS_R (unrestricted) is calculated using the formula;

$$F = \frac{\frac{(RSS_R - RSS_{UR})}{J}}{\frac{RSS_{UR}}{(n-k)}}$$

For this equation J is equals the number of restrictions, *n* is the total of number of observations, and *k* is the number of parameters.

Question 3

Marc Nerlove has estimated the following cost function for electricity generation:

$$Y = AX^{\beta} P^{\alpha_1} P^{\alpha_2} P^{\alpha_3} u \qquad \textbf{(1)}$$

where Y = total cost of production
X = output in kilowatt hours
P_1 = price of labor input
P_2 = price of capital input
P_3 = price of fuel
u = disturbance term

Theoretically, the sum of the price elasticities is expected to be unity, i.e., $(\alpha_1 + \alpha_2 + \alpha_3) = 1$. By imposing this restriction, the preceding cost function can be written as

$$(Y/P_3) = AX^{\beta}(P_1/P_3)^{\alpha_1}(P_2/P_3)^{\alpha_2}u \qquad \textbf{(2)}$$

In other words, (1) is an unrestricted and (2) is the restricted cost function. On the basis of a sample of 29 medium-sized firms, and after logarithmic transformation, Nerlove obtained the following regression results

$$\begin{aligned} \widehat{\ln Y_i} = {} & -4.93 \quad + \ 0.94 \ln X_i + \ 0.31 \ln P_1 \\ \text{se} = {} & (1.96) \qquad (0.11) \qquad (0.23) \\ & -0.26 \ln P_2 + \ 0.44 \ln P_3 \\ & (0.29) \qquad (0.07) \qquad\qquad \text{RSS} = 0.336 \end{aligned} \qquad \textbf{(3)}$$

$$\begin{aligned} \widehat{\ln (Y/P_3)} = {} & -6.55 \ + \ 0.91 \ln X + \ 0.51 \ln (P_1/P_3) + \ 0.09 \ln (P_2/P_3) \\ \text{se} = {} & (0.16) \quad (0.11) \qquad (0.19) \qquad\qquad (0.16) \qquad \text{RSS} = 0.364 \end{aligned} \qquad \textbf{(4)}$$

a. Show the transition from equation (1) to equation (2).

ANSWER:

We begin by taking the log of both sides in equation 1 to derive equation 2.

$$Y = AX^{\beta} P^{\alpha 1} P^{\alpha 2} P^{\alpha 3} u \text{ ------------------------(equation 1)}$$

$$\ln Y = AX^{\beta} (\ln P_1^{\alpha 1})(\ln P_2^{\alpha 2})(\ln P_3^{\alpha 3})u$$

$$\ln Y = AX^{\beta} (\alpha_1 \ln P_1)(\alpha_2 \ln P_2)(\alpha_3 \ln P_3)\, u \text{ --------------------(equation 2)}$$

Next a restriction is imposed on α_3 resulting in,

$$\boldsymbol{\alpha_1 + \alpha_2 + \alpha_3 = 1}$$
$$\boldsymbol{\alpha_3 = 1 - \alpha_1 - \alpha_2}$$

Now we substitute the value obtained for α_3 into equation 2.

$$\ln Y = AX^{\beta} (\alpha_1 \ln P_1)(\alpha_2 \ln P_2)(\ln P_3 - \alpha_1 \ln P_3 - \alpha_2 \ln P_3) - u$$

We then take the common of α_1 & α_2 in order to obtain the following result.

$$\ln Y = AX^{\beta}\, \alpha 1(\ln P1 - \ln P_3)\alpha 2(\ln P2 - \ln P_3)\, u$$

$$\frac{lnY}{lnP_3} = AX^{\beta} (\frac{P_1}{P_3})^{\propto_1} (\frac{P_1}{P_3})^{\propto_2} u$$

b. Interpret Eqs. (3) and (4).

ANSWER:

Both the models discussed above are log-linear, that is taking into consideration the estimated slope coefficients that represent the partial or elasticity of the dependent variable with respect to the regressor that is being taken into consideration. For example, the coefficient of 0.94 in equation 3 basically tells us that if the output in kwh were to increase by 1%, then on average the total cost of production will also increase by 0.94%. When looking at equation 4, the trend described is about the same. If the price of labour comparative to the price of fuel goes up by about 1% then once again on average the relative cost of production also increases by 0.51%.

c. How would you find out if the restriction ($\alpha_1 + \alpha_2 + \alpha_3$) = 1 is valid? Show your calculations.

ANSWER:

In order to find out if the restrictions ($\alpha_1 + \alpha_2 + \alpha_3$) = 1 is valid, the F statistics will be used under the formula, this formula is because the dependent variables in equations 3 and 4 are not the same;

$$F = \frac{\frac{(RSS_R - RSS_{UR})}{NR}}{\frac{(1 - RSS_{UR})}{n-k}}$$

Inputs of the formula includes *n* =29, *k* =5, α =0.05 with the *df* for numerator being 1, and *df* for denominator being 24. Our decision rule for this equation is that if calculated *F* is greater than the critical *F* value then our H_0 is to be rejected.

$$F = \frac{\frac{(RSS_R - RSS_{UR})}{NR}}{\frac{(1 - RSS_{UR})}{n-k}}$$

$$F = \frac{\frac{(0.364 - 0.336)}{1}}{\frac{(1 - 0.336)}{24}}$$

$$\boldsymbol{F = 1.012}$$

The tabulated *F* value is 4.26. Based on the results the *F* obtained is concluded to be insignificant because the 5% critical value of *F* for 1, and 24 numerator and denominator *df* is 4.26. Therefore, we cannot reject the null hypothesis that the sum of the price of elasticities is 1.

Question 4

Suppose theory suggested that annual income (*Y*) depended on gender (*G*), degree received (*D*), and years of experience (*X*): *Y* = *f* (*G*,*D*,*X*). We define four dummy variables:

D_1 = 1 if non-graduate; 0 otherwise
D_2 = 1 if graduate; 0 otherwise

G_1 = 1 if male; 0 otherwise
G_2 = 1 if female; 0 otherwise

First consider the following results obtained from regressing gender dummy with Y:

$$\hat{Y} = \$\ 8000 - 1500\ G_2 \qquad \textbf{(1)}$$
$$\text{se} \qquad (2000) \qquad (600)$$

The sample size is 10.

a. Why is G_1 not included in the equation?

ANSWER:

G_1 is not included in the above equation because it allows us to avoid the trap of a dummy variable or the perfect multi-collinearity. It is possible to include G_1 into the equation but only if we are to drop the intercept or constant.

b. What is the mean income for male and female in the sample?

ANSWER:

Using the equation the mean income for males and females can be found.

Mean Income for males:

$$\hat{Y} = \$\ 8000 - 1500\ G_2$$
$$\hat{Y} = \$\ 8000 - 1500\ (0)$$
$$\hat{Y} = \$\ 8000$$

Mean Income for females

$$\hat{Y} = \$\ 8000 - 1500\ G_2$$
$$\hat{Y} = \$\ 8000 - 1500\ (1)$$
$$\hat{Y} = \$\ 8000 - 1500\ G_2$$
$$\hat{Y} = \$\ 6500$$

c. Do females earn significantly less than males? ($\alpha = 0.05$)

ANSWER:

To being we formulate hypotheses for testing;

H_0: Females earn significantly less than males
H_1: Females earn significantly more than males

Our level of significance is at α=0.05, it is a one tail *t*-test so $df = n - 1$; therefore, 10-1=9, n =9.

The following formula is used;

$$t = \frac{\hat{\beta}_2 - \beta_2}{Se(\beta_2)}$$

The decision rule for this equation will be that if $t_{calculated} > t_{critical}$ then our null hypothesis is rejected.

$$t = \frac{\hat{\beta}_2 - \beta_2}{Se(\beta_2)}$$

$$t = \frac{-1500 - (0)}{600}$$

$$t = -2.5$$

$$t_{critical}(0.05, 9) = 1.833.$$

Based on the results we accept the null hypothesis as $t_{calculated} < t_{critical}$. It is therefore concluded that females earn less than males.

d. What would b_1 and b_2 be if the same sample was used to estimate

$$Y = b_1 + b_2 G_1 + \hat{u}$$

To control for educational attainment and years of experience as well as gender, the following estimates were obtained from the same sample used to estimate equation (1):

$$Y_i = \$ 7000 - 2000 G_{2i} - 6000 D_{2i} + 500 X_i + 1600 G_{2i} D_{2i} + \hat{u}_i$$

ANSWER:

If we substitute G_2 for G_1 in the equation using the same sample the equation will be as follows

$$\hat{Y} = \$6500 + 1500G_1$$

Where b_1 =6500 and b_2 = 1500

e. What is the mean income for a male who is a non-graduate and has no work experience?

ANSWER:

Using the equation

$$Y_i = \$ 7000 - 2000 G_{2i} - 6000 D_{2i} + 500 X_i + 1600 G_{2i} D_{2i} + \hat{u}_i$$

The mean income for a male who is a non-graduate and has no work experience is as follows with variables defined.

For male G_2 =0
For non-graduate D_2 =0
For no work experience X =0

Plugging in the values into the equation we get.

$$\begin{aligned} Y_i &= \$ 7000 - 2000 (0) - 6000 (0) + 500 (0) + 1600 (0) (0) \\ Y_i &= \$ 7000 - 0 - 0 + 0 + 0 \\ Y_i &= \$ 7000 \end{aligned}$$

Hence, mean income for males of this category is \$7000.

f. What is the mean income for a female graduate with 10 years of experience?

ANSWER:

Using the equation,

$$Y_i = \$\ 7000 - 2000\ G_{2i} - 6000\ D_{2i} + 500\ X_i + 1600\ G_{2i}\ D_{2i} + \hat{u}_i$$

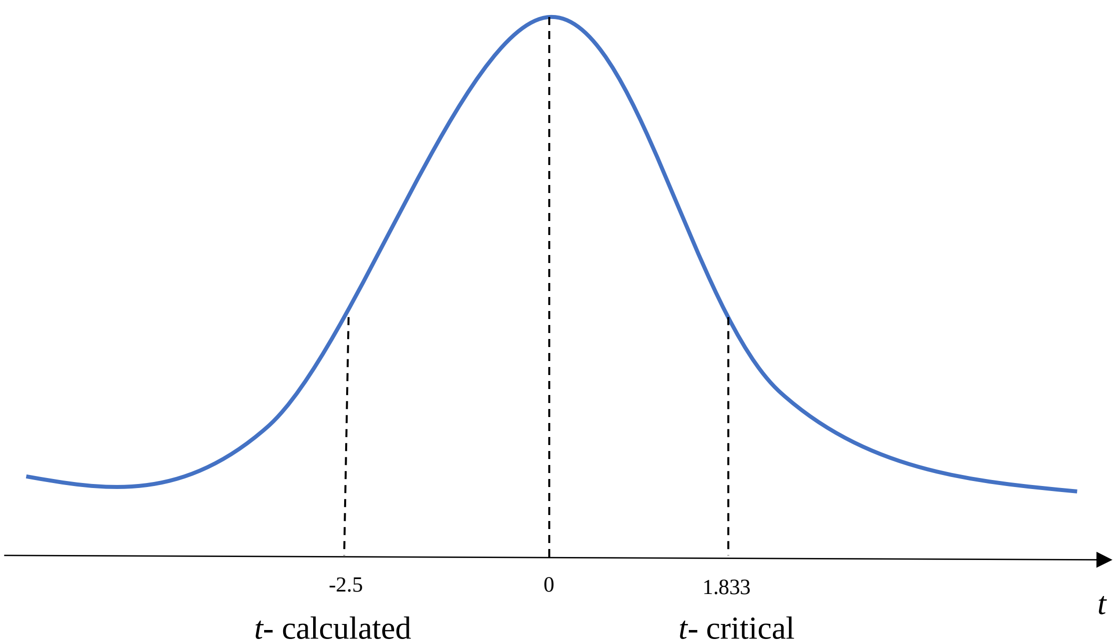
And the following variables,

For female G2 =1
For graduate D2 = 1
For work experience X =10.

Plugging in,

$$Y_i = \$\ 7000 - 2000\ (1) - 6000\ (1) + 500\ (10) + 1600\ (1)(1)$$
$$Y_i = \$\ 7000 - 2000 - 6000 + 5000 + 1600$$
$$Y_i = \$\ 5600$$

Mean income for female graduates who have ten years of experience is $5600.



Question 5

Table 1.1 gives data on the GDP (gross domestic product) deflator for domestic goods and the GDP deflator for imports for country *XYZ* for the period 1968-1982. The GDP deflator is often used as an indicator of inflation in place of CPI.

a. If country *XYZ* is a closed economy, its inflation level can be gauged by GDP deflator for domestic good (Y). How can we estimate the growth rate of GDP deflator (Y) from 1968-82? Estimate the instantaneous growth rate and the growth rate over time [i.e. r in $Y_t = Y_0 (1+r)^t$].

ANSWER:

We begin by taking ln of the equation: $Y_t = Y_0 (1+r)^{t,}$ to estimate the growth rate of GDP reflator Y.

$$
\begin{aligned}
Y_t &= Y_0 (1+r)^t \\
\ln Y_t &= \ln Y_0 + t \ln(1+r) \\
\ln Y_t &= \ln Y_0 + \beta_2 t \\
\ln Y_t &= \beta_1 + \beta_2 t \\
Y &= \$16.0898 + 0.5340 \\
Se &= (40.5031)\ (0.0217) \\
r^2 &= 0.97889672
\end{aligned}
$$

To find instantaneous growth rate;

$$
\begin{aligned}
\beta_2 &= \ln(1 + r) \\
r &= \ln \beta_2 - 1 \\
r &= 1.706 - 1 \\
r &= 0.706 \\
r &= 70.6\%
\end{aligned}
$$

To find growth rate over time;

$$
\begin{aligned}
r &= \beta_2 \text{ x } 100 \\
&= 0.534 \text{ x } 100 \\
&= 53.4\%
\end{aligned}
$$

The rate of growth for durable goods expenditure was about 53.4% but the estimated coefficients are concluded to be indirectly statistically significant. The P values of the set are extremely low meaning it would be inadequate to run a double log model for the circumstance.

$$
\begin{aligned}
\ln Exp &= \beta_1 + \beta_2 \ln \text{time} + u_2 \\
\ln Y_t &= \ln Y_0 + t \ln(1+r) \\
&= \ln \$16.0898 + 1\$ \ (1 + r)
\end{aligned}
$$

b. If country *XYZ* is a small open economy heavily dependent on foreign trade for its survival. To study the relationship between domestic and world prices, you are given the following linear model:

$$Y_t = \beta_1 + \beta_2 X_t + u_t$$

i. What are your prior expectations about the relationship between Y and X?

ANSWER:

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.989392096
R Square	0.97889672
Adjusted R Square	0.97727339
Standard Error	52.13170284
Observations	15

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	1638830.646	1638831	603.0179701	2.81131E-12
Residual	13	35330.28773	2717.714		
Total	14	1674160.933			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	516.0898305	40.5631118	12.72313	1.03265E-08	428.4585552	603.7211059	428.4585552	603.7211059
X Variable 1	0.533969258	0.021744585	24.55642	2.81131E-12	0.486992938	0.580945579	0.486992938	0.580945579

ii. Estimate β_1 and β_2. Do your results comply with your prior expectations? What is the economic meaning of the parameter estimates?

ANSWER:

The following table represent the calculations conducted to establish the values of β_1 and β_2.

Year	Y	X	yi=y-y bar	xi=x-xbar	xi sq	xiyi	XY	X²	Yhat	μ=Y-Y hat	Yhat*μ	Y=Yhat +μ	μ²	yi ²	Ŷ-Ybar	ESS
1968	1000	1000	-455.733	-759.733	577194.7	346235.8044	346235.8044	577194.7	1050.0591	-50.0591	-52565	1000	2505.912	207692.9	-20785.9	432055339.6
1969	1023	1042	-432.733	-717.733	515141.1	310587.1378	310587.1378	515141.1	1072.4858	-49.4858	-53072.8	1023	2448.844	187258.1	-20763.5	431123522
1970	1040	1092	-415.733	-667.733	445867.8	277599.0044	277599.0044	445867.8	1099.1843	-59.1843	-65054.4	1040	3502.777	172834.2	-20736.8	430015527
1971	1087	1105	-368.733	-654.733	428675.7	241422.0044	241422.0044	428675.7	1106.1259	-19.1259	-21155.6	1087	365.7986	135964.3	-20729.9	429727681.8
1972	1146	1110	-309.733	-649.733	422153.4	201244.0711	201244.0711	422153.4	1108.7957	37.20429	41251.96	1146	1384.159	95934.74	-20727.2	429616997.8
1973	1285	1257	-170.733	-502.733	252740.8	85833.33778	85833.33778	252740.8	1187.2892	97.71081	116011	1285	9547.403	29149.87	-20648.7	426369258.2
1974	1485	1749	29.267	-10.7333	115.2044	-314.1288889	-314.1288889	115.2044	1450.0021	34.99794	50747.08	1485	1224.856	856.5378	-20386	415588911.9
1975	1521	1770	65.267	10.26667	105.4044	670.0711111	670.0711111	105.4044	1461.2154	59.78458	87358.15	1521	3574.196	4259.738	-20374.8	415131846.8
1976	1543	1889	87.267	129.2667	16709.87	11280.67111	11280.67111	16709.87	1524.7578	18.24224	27815	1543	332.7793	7615.471	-20311.2	412546561.4
1977	1567	1974	111.267	214.2667	45910.2	23840.73778	23840.73778	45910.2	1570.1451	-3.14515	-4938.34	1567	9.891946	12380.27	-20265.9	410704872.9
1978	1592	2015	136.267	255.2667	65161.07	34784.33778	34784.33778	65161.07	1592.0379	-0.03789	-60.316	1592	0.001435	18568.6	-20244	409818002.1
1979	1714	2260	258.267	500.2667	250266.7	129202.2044	129202.2044	250266.7	1722.8604	-8.86035	-15265.2	1714	78.50588	66701.67	-20113.1	404538386.4

1980	1841	2621	385.267	861.2667	741780.3	331817.3378	331817.3378	741780.3	1915.6233	-74.6233	-142950	1841	5568.63	148430.4	-19920.4	396821409.6
1981	1959	2777	503.267	1017.267	1034831	511956.4044	511956.4044	1034831	1998.9225	-39.9225	-79801.9	1959	1593.803	253277.3	-19837.1	393509645.3
1982	2033	2735	577.267	975.2667	951145.1	562988.9378	562988.9378	951145.1	1976.4958	56.50425	111680.4	2033	3192.73	333236.8	-19859.5	394399909
SUMS	21836	26396			**5747799**	**3069147.933**				**0**	**0**	**21836**	**35330.29**	**1674161**		**6231967872**
MEANS	1455.733	1759.733											**RSS**	**TSS**		

Finding β2:

$$\beta_2 = \frac{x_i y_i}{x_i^2}$$

$$\beta_2 = \frac{3069147.933}{5747799}$$

$$\beta_2 = 0.533969$$

Finding β1:

$$\beta_1 = Mean\ of\ Y - \beta_2 \times Mean\ of\ X$$
$$\beta_1 = 1455.733 - 0.533969 \times 1759.733$$
$$\beta_1 = 516.0898$$

iii. Also estimate the model using the double log functional form. Provide the economic meaning of the parameter estimates?

ANSWER:

lnY	lnX
3	3
3.009875634	3.01786772
3.017033339	3.03822264
3.036229544	3.04336228
3.059184618	3.04532298
3.108903128	3.09933528
3.171726454	3.24278981
3.182129214	3.24797327
3.188365926	3.27623196
3.195068996	3.29534715
3.201943063	3.30427505
3.234010818	3.35410844
3.265053789	3.41846702
3.292034436	3.44357588
3.308137379	3.43695733

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.990685216
R Square	0.981457197
Adjusted R Square	0.980030828
Standard Error	0.014909774
Observations	15

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	0.152961262	0.152961	688.0806	1.21139E-12
Residual	13	0.002889918	0.000222		
Total	14	0.15585118			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	1.08295234	0.078944813	13.71784	4.14E-09	0.912402441	1.25350224	0.912402441	1.25350224
X Variable 1	0.642829358	0.024506203	26.23129	1.21E-12	0.589886924	0.695771792	0.589886924	0.695771792

lnY	X
3	1000
3.009875634	1042
3.017033339	1092
3.036229544	1105
3.059184618	1110
3.108903128	1257
3.171726454	1749
3.182129214	1770
3.188365926	1889
3.195068996	1974
3.201943063	2015
3.234010818	2260
3.265053789	2621
3.292034436	2777
3.308137379	2735

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.977612883

R Square	0.955726949
Adjusted R Square	0.95232133
Standard Error	0.023038441
Observations	15

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	0.148951173	0.148951	280.6324	3.50778E-10
Residual	13	0.006900007	0.000531		
Total	14	0.15585118			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	2.868031702	0.017925961	159.9932	8.37E-23	2.829305017	2.906758386	2.829305017	2.906758386
X Variable 1	0.00016098	9.60953E-06	16.75208	3.51E-10	0.00014022	0.00018174	0.00014022	0.00018174