2020

STATISTICS — HONOURS

Paper : DSE-A-2

(Econometrics)

Full Marks : 50

The figures in the margin indicate full marks.

Candidates are required to give their answers in their own words as far as practicable.

1. Answer ***any ten*** questions : 1×10

(a) Distinguish between an *economic* model and an *econometric* model.

(b) The expenditure (E) of a household on a particular commodity depends on the income (I) of the household and the price level (P). An economic model is set up as follows :

$$E = \alpha I^{\beta} P^{\gamma}$$

Give economic interpretations of the parameters β and γ.

(c) When is it necessary to transform the *response* variable?

(d) If in a three-regressor model, the response y is linear with respect to x_1 and x_2, but has a non-linear relationship with x_3, how will you set up your model?

(e) A salesman receives a commission on the number of units sold (x) of a commodity. The rate of commission is β_1 for the first 500 units, β_2 ($> \beta_1$) for the next 1000 units and $\beta_3 (> \beta_2)$ for all units above 1000. Set up a model to show how the saleman's total commission (y) changes with units sold.

(f) For a regression model with a continuous response variable and a single binary regressor (x), the regression coefficients are (intercept = 5, x-coefficient = 2). Give a rough sketch of the predicted response against the regressor.

(g) For the linear model, $y_i = \beta_0 + \beta_1 x_{1i} + ... + \beta_p x_{pi} + u_i$, $i = 1, ..., n$, state the *necessary* conditions under which the ordinary least squares estimators of the regression parameters are *best linear unbiased*.

(h) What effect will the use of *ordinary least squares* in the presence of heteroscedasticity have in interval estimation of regression parameters?

(i) While performing a Glejser's test for checking heteroscedasticity, the intercept comes out to be significant, but the slope (of the single regressor) is insignificant. What would be your inference?

(j) The problem of autocorrelation is discussed only in terms of time indexed data. Why is the problem not addressed for cross-sectional data?

(k) What would be your conclusion if the Durbin-Watson statistic comes out to be 3.8?

(l) Given three regressors which exhibit multicollinearity, how will you decide which variable(s) to drop to get rid of the problem?

(m) What is the Variance Inflation Factor (VIF)?

(n) What problem do measurement error of response variables create?

(o) What are *instrumental variables*?

2. Answer ***any four*** questions : 5×4

(a) Set up a regression model with income as response and years of experience (continuous), gender (male/female) and educational qualification (Madhyamik/HS/Graduate/Post graduate) as regressors, with gender also affecting the *increment* in income with years of experience.

(b) For the linear regression model, $y_i = \beta_0 + \beta_1 x_{1i} + ... + \beta_p x_{pi} + u_i,\ i = 1, ..., n,$ with u_i identically and independently distributed with mean zero and variance σ^2, find an unbiased estimator of σ^2.

(c) For a p-variable model with error terms having a non-spherical distribution, show that the generalized least squares give *better* estimators than the ordinary least squares.

(d) How will you estimate the regression parameters of a first-order autocorrelated error model when the autoregressive parameter is unknown?

(e) Why is the ordinary least squares method not used for estimating the regression parameters in the presence of multicollinearity?

(f) Justify the use of Bartlett's method in estimating the parameters of a single-regressor model when the regressor is measured with error. Derive the estimators.

3. Answer ***any two*** questions : 10×2

(a) In the linear regression model,

$$y_i = \beta_0 + \beta_1 x_{1i} + ... + \beta_p x_{pi} + u_i,\ i = 1, ..., n,$$

Where the u_is are independently distributed with zero mean, it is conjectured that $V(u_i)$ is proportional to the inverse of $\alpha_0 + \alpha_1 x_{1i}$.

(i) Describe Goldfeld-Quandt's technique to test this conjecture.

(ii) If the conjecture is correct, describe how you will estimate $\beta_0, \beta_1, ..., \beta_p$.

(b) Consider the model,

$$y_t = \beta_0 + \beta_1 x_t + u_t,$$
$$u_t = \rho u_{t-1} + \varepsilon_t, \qquad t = 1, ..., n,$$

where ε_t are identically and independently distributed with mean zero and variance σ^2.

Write $u = (u_1, u_2, ..., u_n)'$.

(i) Find $E(u)$ and the error dispersion matrix $D(u)$.

(ii) To test $H_0 : \rho = 0$, why cannot a simple t-test, based on the autoregressive equation, be used? Can we use the residuals instead? Or else, how will you perform the test?

(c) (i) Why is multicollinearity referred to as a *sample phenomenon*? Explain in this context why getting more data often helps to overcome multicollinerity.

(ii) What are the consequences of the presence of multicollinearity? Justify the use of Condition Numbers to detect it.

———————