ECO601 Business Econometrics — 100 Conceptual MCQs

Q1. Econometrics combines which three fields?

A. History, Geography, Physics
B. Economics, Mathematics, Statistics
C. Finance, Law, Accounting
D. Biology, Chemistry, Math

Q2. Which software is commonly used for econometric analysis at Virtual University?

A. AutoCAD
B. Adobe Photoshop
C. Microsoft Excel and Stata
D. MATLAB only

Q3. OLS stands for:

A. Ordered Least Sum
B. Ordinary Least Squares
C. Optional Linear System
D. Observed Log Statistics

Q4. In a regression Y = a + bX + e, the letter 'e' represents:

A. Euler's number
B. The slope
C. The error or residual term
D. The intercept

Q5. The coefficient of determination R^2 measures:

A. The slope of the regression line
B. The proportion of variation in Y explained by X
C. The number of observations
D. The standard deviation of errors

Q6. If $R^2 = 0.85$, it means the model explains:

A. 15% of variation in Y
B. 85% of variation in Y
C. 0.85 units of Y
D. 85 observations

Q7. Homoskedasticity means the error term has:

A. Increasing variance
B. Decreasing variance
C. Constant variance across all observations
D. Zero variance

Q8. Heteroskedasticity is more common in:

A. Experimental lab data
B. Cross-sectional data
C. Data with no variables
D. Perfectly measured data

Q9. Multicollinearity refers to:

A. Errors correlated with Y

B. High linear relationship among independent variables
C. Non-linear relationship between X and Y
D. Missing values in the dataset

Q10. VIF stands for:

A. Variable Importance Factor
B. Variance Inflation Factor
C. Value in Forecasting
D. Variable Interaction Function

Q11. A VIF value greater than 10 indicates:

A. No multicollinearity
B. Mild multicollinearity
C. Problematic multicollinearity
D. Heteroskedasticity

Q12. The t-test in regression is used to test:

A. Overall model fit
B. Significance of individual coefficients
C. Presence of multicollinearity
D. Normality of residuals

Q13. The F-test in regression is used to test:

A. Individual variable significance
B. Overall goodness of fit of the model
C. Heteroskedasticity
D. Autocorrelation

Q14. A dummy variable takes values:

A. Any real number
B. Only 0 and 1
C. Only positive integers
D. Between -1 and 1

Q15. For a qualitative variable with 4 categories, how many dummies are needed in regression?

A. 4
B. 5
C. 3
D. 2

Q16. The Dummy Variable Trap occurs when:

A. Too few dummies are used
B. All category dummies are included with an intercept causing perfect multicollinearity
C. Dummies are used for continuous variables
D. The intercept is removed

Q17. Cross-sectional data is collected:

A. Over many time periods
B. From the same units followed over time
C. From different entities at a single point in time
D. Only from government sources

Q18. Panel data is also called:

A. Pooled data
B. Cross-sectional data
C. Longitudinal data
D. Time series only

Q19. In the double-log model lnY = a + b lnX, the coefficient b gives:

A. The marginal effect in units
B. The elasticity of Y with respect to X
C. The intercept
D. The standard error

Q20. The log-level model lnY = a + bX implies that a one unit increase in X changes Y by approximately:

A. b units
B. b percent
C. b × 100 percent
D. 1/b percent

Q21. Adjusted R^2 differs from R^2 because it:

A. Is always higher than R^2
B. Penalises for adding unnecessary variables
C. Equals R^2 in all cases
D. Is used only in simple regression

Q22. The Goldfeld-Quandt test is used to detect:

A. Multicollinearity
B. Autocorrelation
C. Heteroskedasticity
D. Missing data

Q23. The White test is preferred for detecting heteroskedasticity because:

A. It is the oldest test
B. It requires the smallest sample
C. It is general and does not require normality assumption
D. It always finds heteroskedasticity

Q24. The Breusch-Pagan test regresses _____ on the original X variables:

A. Y
B. Squared residuals e^2
C. Absolute residuals |e|
D. Log of Y

Q25. BLUE stands for:

A. Best Linear Unbiased Estimator
B. Basic Least Unit Estimator
C. Binary Logarithmic Unbiased Equation
D. Biased Linear Uniform Estimator

Q26. An unbiased estimator means:

A. It always gives the exact true value
B. Its average value over repeated samples equals the true parameter
C. It has zero variance

D. It is always positive

Q27. Ceteris paribus means:

A. All variables change together
B. Other things being equal (held constant)
C. Variables are randomly assigned
D. The model has no intercept

Q28. In the regression equation Y = a + bX, increasing X by one unit changes Y by:

A. a units
B. b units on average
C. 1 unit always
D. X units

Q29. The intercept 'a' in Y = a + bX represents:

A. The slope
B. The value of Y when X = 0
C. The mean of Y
D. The residual

Q30. Price elasticity of demand in a linear model Qd = a + bP equals:

A. b
B. b × (P/Q)
C. 1/b
D. a/b

Q31. In the Cobb-Douglas production function $Y = AL^{\alpha}K^{\beta}$, taking natural log gives:

A. $Y = A + \alpha L + \beta K$
B. $\ln Y = \ln A + \alpha \ln L + \beta \ln K$
C. $\ln Y = \ln A \times \ln L \times \ln K$
D. $Y = \ln A + \alpha + \beta$

Q32. Which distribution is used for the goodness of fit test in regression?

A. Normal distribution
B. t-distribution
C. F-distribution
D. Uniform distribution

Q33. Degrees of freedom in simple regression with n=20 observations is:

A. 20
B. 18
C. 19
D. 2

Q34. Omitting an important variable from a regression model causes:

A. Heteroskedasticity only
B. Omitted variable bias in estimated coefficients
C. Perfect multicollinearity
D. R^2 to equal 1

Q35. A scatter diagram is used before regression to:

A. Estimate coefficients
B. Identify the likely functional form of the relationship

C. Test for normality
D. Compute R^2

Q36. The normal distribution is symmetric around its:

A. Variance
B. Standard deviation
C. Mean
D. Mode only when skewed

Q37. The t-distribution approaches the normal distribution as:

A. Sample size decreases
B. Variance increases
C. Degrees of freedom increase toward infinity
D. The intercept decreases

Q38. Mutually exclusive events mean:

A. Both can occur at the same time
B. One event prevents the other from occurring simultaneously
C. Both are equally likely
D. One event causes the other

Q39. Classical probability of an event A equals:

A. Number of unfavourable outcomes / Total outcomes
B. Number of favourable outcomes / Total outcomes
C. Total outcomes / Sample space
D. 1 minus P(B)

Q40. A continuous random variable can take:

A. Only whole number values
B. A finite countable set of values
C. Any value within an interval
D. Only values 0 and 1

Q41. The Chi-square distribution is related to the normal distribution because:

A. They have the same mean
B. The square of a standard normal variable follows a Chi-square distribution
C. Both are symmetric
D. Chi-square is a subset of normal

Q42. The F-distribution is a ratio of:

A. Two normal distributions
B. Two t-distributions
C. Two Chi-square distributions divided by their degrees of freedom
D. Mean and variance

Q43. Listwise deletion in handling missing data means:

A. Replacing missing values with the mean
B. Removing all rows that contain at least one missing value
C. Using separate analyses for each variable
D. Running regression multiple times

Q44. MCAR stands for:

A. Missing Completely at Random

B. Missing Cannot be Analyzed in Regression
C. Multiple Correlation Adjustment Required
D. Maximum Conditional Autoregressive Residual

Q45. In multiple imputation, the main advantage over single imputation is:

A. It is faster
B. It generates multiple datasets preserving variability and reducing bias
C. It always gives the exact true value
D. It requires no model

Q46. An interaction term X1 × X2 in regression captures:

A. The sum of effects of X1 and X2
B. How the effect of X1 on Y changes depending on the value of X2
C. The correlation between X1 and X2
D. The squared effect of X1

Q47. Standardised regression coefficients (Beta coefficients) allow:

A. Comparison of effect sizes across variables with different units
B. Prediction of exact Y values
C. Removal of the intercept
D. Testing for heteroskedasticity

Q48. After standardising variables, the intercept of the regression is always:

A. Equal to the mean of Y
B. Zero
C. Equal to $\beta 1$
D. Undefined

Q49. A quadratic regression $Y = a + bX + cX^2$ is still estimated by OLS because:

A. X^2 is a new variable and the model is linear in parameters
B. The model is linear in both variables and parameters
C. OLS cannot handle squared terms
D. The quadratic term eliminates the error

Q50. An inverted-U shaped relationship between two variables can be modelled using:

A. A linear equation
B. A quadratic equation with negative coefficient on X^2
C. A log-log model
D. A dummy variable

Q51. Tolerance = 1 - R^2 from an auxiliary regression. A Tolerance of 0.05 means:

A. No multicollinearity
B. 95% of the variable is explained by other regressors — severe multicollinearity
C. Mild multicollinearity
D. R^2 of the main regression is 0.05

Q52. The remedy of 'first differencing' (Yt - Yt-1) for multicollinearity results in:

A. Gaining one extra observation
B. Losing one observation
C. Doubling the sample size
D. Making all coefficients zero

Q53. Adding more independent variables to a regression always:

A. Decreases R^2
B. Increases R^2 (it cannot decrease)
C. Has no effect on R^2
D. Decreases adjusted R^2

Q54. The Park LM test auxiliary regression uses:

A. e^2 regressed on X
B. $\ln e^2$ regressed on lnX
C. $|e|$ regressed on X
D. Y regressed on e^2

Q55. The Glesjer test uses which dependent variable in the auxiliary regression?

A. e^2
B. $\ln e^2$
C. |e| (absolute value of residuals)
D. Y

Q56. In the LM test, the test statistic is computed as:

A. N / R^2
B. $N \times R^2$ (from the auxiliary regression)
C. R^2 / N
D. $(N-k) \times R^2$

Q57. The Goldfeld-Quandt test splits the data into groups after:

A. Sorting by Y values
B. Sorting by the suspected X variable and omitting middle observations
C. Sorting by residuals
D. Random splitting

Q58. Under heteroskedasticity, OLS estimators are:

A. Biased and inconsistent
B. Unbiased and consistent but inefficient
C. Efficient and biased
D. Neither unbiased nor consistent

Q59. A lagged variable model Yt = a + bXt-1 + e is used when:

A. X has an immediate effect on Y
B. The effect of X on Y occurs with a time delay
C. Y causes X
D. The model is non-linear

Q60. The Kuznets Curve describes an inverted-U relationship between:

A. Price and quantity
B. Income per capita and income inequality
C. Capital and labour
D. Exports and imports

Q61. A regression residual is defined as:

A. Yi - X■
B. Yi - ■i (actual minus fitted value)
C. ■i - ■
D. Yi - β0

Q62. OLS minimises the sum of squared residuals rather than the sum of residuals because:

A. Squared residuals are always smaller
B. Positive and negative residuals cancel out; squaring magnifies and avoids cancellation
C. It is computationally easier
D. The sum of residuals is always zero anyway

Q63. In a semi-log model lnY = a + bX, if b = 0.05, a one unit increase in X increases Y by approximately:

A. 0.05 units
B. 5 percent
C. 50 percent
D. 0.5 units

Q64. Which of the following is a time series example?

A. GDP of 50 countries in 2020
B. Wages of 200 workers surveyed in one month
C. Pakistan's inflation rate from 2000 to 2023
D. Sales of 100 companies in 2022

Q65. The formula for the slope b in simple OLS regression is:

A. $\Sigma XY / \Sigma X^2$
B. $(n\Sigma XY - \Sigma X\Sigma Y) / (n\Sigma X^2 - (\Sigma X)^2)$
C. $\Sigma Y / n - a$
D. $\Sigma X / \Sigma Y$

Q66. When the F-statistic is significant in regression, it means:

A. All individual coefficients are significant
B. At least one slope coefficient is significantly different from zero
C. R^2 equals 1
D. There is no heteroskedasticity

Q67. The standard error of the estimate measures:

A. The slope of the regression
B. The accuracy of predictions — standard deviation of the residuals
C. The coefficient of determination
D. The number of outliers

Q68. P-value in regression output represents:

A. The value of the parameter
B. The probability of observing a test statistic as extreme as computed if H0 is true
C. The coefficient of determination
D. The standard error

Q69. If P-value = 0.002 and α = 0.05, we:

A. Fail to reject H0
B. Reject H0 — the coefficient is significant
C. Cannot conclude anything
D. Accept H0

Q70. The reference category in dummy variable regression is the group that:

A. Has the highest mean
B. Has all dummy variables set to zero — the baseline comparison group

C. Is always the majority group
D. Is excluded from the analysis

Q71. Which test checks if variance of errors increases or decreases with an explanatory variable by comparing two sub-sample RSS values?

A. White test
B. Breusch-Pagan test
C. Goldfeld-Quandt test
D. Park test

Q72. In regression analysis, a high R^2 with individually insignificant t-statistics most likely indicates:

A. Heteroskedasticity
B. Multicollinearity
C. Autocorrelation
D. Model is perfectly specified

Q73. The Cobb-Douglas function exhibits constant returns to scale when:

A. $\alpha + \beta > 1$
B. $\alpha + \beta = 1$
C. $\alpha + \beta < 1$
D. $\alpha = \beta$

Q74. Pooled cross-sectional data differs from panel data because in pooled data:

A. The same individuals are tracked each period
B. Different individuals may appear in different time periods
C. There is only one time period
D. Only time series variation is used

Q75. An auxiliary regression in multicollinearity detection regresses:

A. Y on all X variables
B. One X variable on all other X variables
C. Residuals on Y
D. Y on its own lagged values

Q76. The normal equations in OLS are derived by setting partial derivatives of $\Sigma\Sigma e^2$ equal to:

A. One
B. The slope
C. Zero
D. R^2

Q77. An econometric model is said to be 'well-specified' when:

A. R^2 is maximised by including all possible variables
B. The model includes the correct variables in the correct functional form based on theory
C. The intercept is zero
D. All coefficients are positive

Q78. In a wage regression, if the female dummy coefficient is -20 and significant, it means:

A. Women have 20 more years of experience
B. Women earn on average 20 units less than men with the same education and experience
C. Women are 20 percent more productive
D. The model excludes men

Q79. The power rule of differentiation states that if y = ax■ then dy/dx =

A. ax■
B. anx^(n-1)
C. ax^(n+1)
D. n/x

Q80. A first order condition for optimization requires setting the _____ equal to zero:

A. Second derivative
B. First derivative
C. Function itself
D. Integral

Q81. A second order condition confirms a maximum when the second derivative is:

A. Positive
B. Zero
C. Negative
D. Undefined

Q82. The product rule states that if y = f(x)·g(x), then dy/dx =

A. f'(x)·g'(x)
B. g(x)·f'(x) + f(x)·g'(x)
C. f(x)/g'(x)
D. f'(x) + g'(x)

Q83. Sample space is defined as:

A. The probability of a single event
B. The set of all possible outcomes of an experiment
C. A subset of outcomes
D. The mean of the distribution

Q84. Which summation property allows a constant to be factored out: Σ(kXi) =

A. k + ΣXi
B. k × ΣXi
C. ΣXi / k
D. k² × ΣXi

Q85. The partial derivative ∂Z/∂x means we differentiate Z with respect to x while:

A. All variables including y change
B. Treating all other variables as constants
C. Setting Z equal to zero
D. Differentiating twice

Q86. Cross partial derivatives Zxy and Zyx are always:

A. Opposite in sign
B. Equal to each other
C. Both equal to zero
D. Both negative

Q87. In the Keynesian consumption function C = a + bY, the coefficient b is called:

A. Autonomous consumption
B. Marginal propensity to consume
C. National income

D. Investment multiplier

Q88. Econometric forecasting is most reliable when:

A. The model is incorrectly specified

B. The model is well-specified and estimated appropriately with diagnostic tests passed

C. R^2 is exactly 0.5

D. Only one variable is used

Q89. Which of the following is NOT a step in the econometric methodology?

A. Model Specification

B. Data Collection

C. Publishing raw data without analysis

D. Hypothesis Testing

Q90. Stata software is introduced in ECO601 because:

A. It is free and simple

B. It is more powerful than Excel for complex econometric models and provides robust standard errors

C. It replaces mathematics entirely

D. It is only used for graphics

Q91. The Analysis ToolPak in Excel is limited to a maximum of _____ independent variables:

A. 10

B. 50

C. 16

D. 5

Q92. A confidence interval for a regression coefficient that does NOT include zero means:

A. The coefficient is not significant

B. The coefficient is statistically significant at that confidence level

C. The model is wrong

D. R^2 is zero

Q93. The concept of 'empirical content' in econometrics means:

A. Using only theoretical models without data

B. Testing economic theories against real-world data

C. Collecting data without any theory

D. Using qualitative descriptions only

Q94. An outlier that drastically changes the slope when removed is called:

A. A leverage point

B. An influential point

C. A missing value

D. A residual

Q95. If Tolerance = 0.9 for a variable, there is:

A. Severe multicollinearity

B. Mild multicollinearity

C. No significant multicollinearity

D. Perfect collinearity

Q96. When we say a variable is 'significant at the 5% level', it means:

A. The variable explains 5% of variation

B. We reject H0 at a 5% probability of Type I error

C. The t-statistic equals 5

D. The p-value is 0.5

Q97. A saddle point in multivariate optimisation occurs when:

A. Both second-order partial derivatives are positive

B. Both second-order partial derivatives are negative

C. The second-order partials have opposite signs

D. The first-order conditions are not met

Q98. Which data source is mentioned in ECO601 for world macroeconomic data?

A. Facebook

B. World Development Indicators (World Bank)

C. Wikipedia

D. Google Scholar only

Q99. The sum of OLS residuals Σei always equals:

A. R^2

B. 1

C. 0

D. n

Q100. Econometrics is described as 'making data speak' because:

A. Data automatically generates conclusions

B. Without econometric tools, raw data cannot reveal relationships or test theories

C. Data speaks in multiple languages

D. Computers replace economists