Business Econometrics — ECO601

Multiple Choice Questions Bank

Lectures 1 – 18 | Virtual University of Pakistan

Total Questions: 500 | Topics Covered: Lectures 1–18 | Correct answers are highlighted in green

Topic Coverage by Lecture

Lecture	Topic	Questions
Lecture 1	Overview of Econometrics & Data Types	26
Lecture 2	Summation Algebra & Linear Functions	23
Lecture 3	Quadratic, Exponential, Logarithmic Functions & Derivatives	25
Lecture 4	Partial Derivatives & Simple Optimization	21
Lecture 5	Multivariate Optimization & Probability Review	36
Lecture 6	Simple Regression Model & OLS	39
Lecture 7	Estimation, Testing & R^2 in Regression	45
Lecture 8	Excel Applications in Regression	21
Lecture 9	Multiple Regression	29
Lecture 10	K-Variable Regression & Analysis ToolPak	25
Lecture 11	Transformation & Functional Forms in Regression	30
Lecture 12	Standardised Variables & Regression Strategy	22
Lecture 13	Qualitative Variables & Dummy Variables	26
Lecture 14	Logarithms in Regression & Missing Data	21
Lecture 15	Multicollinearity	33
Lecture 16	Multicollinearity: Remedial Measures	21
Lecture 17	Heteroskedasticity: Causes & Detection	24
Lecture 18	Heteroskedasticity: Formal Tests	33

Lecture 1: Overview of Econometrics & Data Types

Q1. Econometrics is best described as the intersection of which three disciplines?

A. Economics, Physics, and Computer Science
✓ **B. Economics, Mathematics, and Statistics**
C. Finance, Accounting, and Mathematics
D. Sociology, Psychology, and Statistics

Q2. Which of the following is a traditional definition of Econometrics?

A. The use of qualitative methods to study economic behaviour
✓ **B. A branch concerned with the use of mathematical methods (especially statistics) in describing economic systems**
C. A purely theoretical framework for building economic models
D. A branch of accounting focused on financial ratios

Q3. According to the modern view, Econometrics is used in which additional fields beyond Economics?

A. Physics and Chemistry only
✓ **B. Finance, Marketing, Management, and Sociology**
C. Medicine and Biology
D. Law and Political Science

Q4. Which software is mentioned in ECO601 for use BEFORE mid-term examination?

A. SPSS
B. Stata
✓ **C. Microsoft Excel**
D. SAS

Q5. The Economist's Dictionary of Economics defines Econometrics as setting up mathematical models describing economic relationships and then doing what?

A. Making predictions only
✓ **B. Testing validity of hypotheses and estimating parameters**
C. Collecting raw data
D. Publishing academic journals

Q6. Which step comes FIRST in the methodology of Business Econometrics?

A. Model Estimation
B. Collection of Data
✓ **C. Creation of a statement of theory or hypothesis**
D. Prediction or Forecasting

Q7. Cross-Sectional Data is defined as:

A. Observations of the same entity over time
✓ **B. A sample of entities at a given point in time**
C. Combined cross-sections from different years
D. Time series of each cross section

Q8. Panel (Longitudinal) Data differs from Pooled Cross-Sectional Data because in Panel Data:

A. Individual entities change every year
✓ **B. Same cross-sectional units are followed over time**
C. Only one observation per entity is recorded
D. Data is collected from different countries

Q9. Which of the following is an example of Time Series Data?

A. Monthly income of a sample of individuals in 2014
B. GDP across 50 countries in a single year
✓ **C. Pakistan's GDP from 1972 to 2012**
D. Annual sales of 100 companies in 2014

Q10. Which file extension is NOT listed in the lecture as a common data file type?

A. .xls or .xlsx
B. .sav
✓ **C. .pdf**
D. .csv

Q11. Which prescribed textbook author wrote 'Introductory Econometrics: A Modern Approach'?

A. Gujarati
✓ **B. Wooldridge**
C. Green
D. Butt

Q12. Performing 'Diagnostic Tests' in the econometric methodology is done to:

A. Finalise the data collection
✓ **B. Check the validity of the model and diagnose possible estimation problems**
C. Set up the initial theory
D. Publish the results

Q13. The phrase 'Data never speak for themselves; Econometrics makes Data speak' captures which benefit of Econometrics?

A. It replaces all forms of theory
✓ **B. It provides empirical evidence for theoretical statements**
C. It eliminates the need for data collection
D. It makes qualitative research unnecessary

Q14. Which website of the World Bank Group is mentioned as a data archive for macroeconomic and financial data?

A. http://sbp.org.pk
B. http://www.ats.ucla.edu/stat/stata/
✓ **C. http://data.worldbank.org/data-catalog/world-development-indicators**
D. http://finance.gov.pk

Q15. Pooled Cross-Sectional Data is characterized by:

A. The same individuals being surveyed every year
✓ **B. Combining cross-sections from different years where individuals may change**
C. Time series observations on a single individual
D. Only two time periods of data

Lecture 2: Summation Algebra & Linear Functions

Q16. The Greek letter used to indicate summation or addition is:

A. Δ (Delta)
✓ **B. $\sum$ (Sigma)**
C. β (Beta)
D. π (Pi)

Q17. Which property of summation states that the sum of a constant 'a' added n times equals n×a?

A. Property 2
B. Property 3
✓ C. Property 1
D. Property 4

Q18. According to summation Property 2, which of the following is TRUE?

A. $\sum kXi = k + \sum Xi$
✓ B. $\sum kXi = k \times \sum Xi$
C. $\sum kXi = (\sum Xi)/k$
D. $\sum kXi = \sum Xi - k$

Q19. Which of the following is NOT a valid operation in summation algebra?

A. $\sum(Xi + Yi) = \sum Xi + \sum Yi$
✓ B. $\sum(XiYi) = (\sum Xi)(\sum Yi)$
C. $\sum kXi = k\sum Xi$
D. $\sum a = na$ (where a is a constant)

Q20. In a linear function Y = a + bX, the term 'a' represents the:

A. Slope of the line
B. Rate of change
✓ C. Y-intercept
D. Marginal effect of X

Q21. In the linear function Y = a + bX, the slope 'b' is equal to:

A. $\Delta X/\Delta Y$
✓ B. $\Delta Y/\Delta X$
C. Y/X
D. a/Y

Q22. For the simple demand function Qd = a + bP, what is the expected sign of 'b' for a normal good?

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

Q23. The double log functional form lnY = β0 + β1 lnX gives an elasticity equal to:

A. β0
B. β1(X/Y)
✓ C. β1
D. 1/β1

Q24. In the log-linear form lnY = β0 + β1X, what does a one-unit change in X imply for Y?

A. Y changes by β1 units
✓ B. Y changes by approximately β1 × 100 percent
C. Y changes by β1/X units
D. Y remains unchanged

Q25. The Cobb-Douglas production function $Y = AL^{\alpha}K^{\beta}$ can be linearized by:

A. Squaring both sides
✓ B. Taking the natural log of both sides

C. Dividing by L
D. Differentiating with respect to K

Q26. ∑(Xi/Yi) is equal to:

A. ∑Xi / ∑Yi
✓ **B. Neither ∑Xi / ∑Yi nor any simple expression**
C. ∑Xi × ∑Yi
D. n × (Xi/Yi)

Q27. Which of the following correctly expresses ∑(X − X■)(Y − ■)?

A. ∑XY + nX■■
✓ **B. ∑XY − (∑X∑Y)/n**
C. ∑XY × n
D. ∑X × ∑Y / n

Q28. A linear function is defined as one where:

A. The variable X has a power greater than one
✓ **B. The graph forms a straight line and the rate of change is constant**
C. The function must pass through the origin
D. The slope changes at every point

Q29. Double summation notation ∑∑ Xij represents:

A. Squaring the single summation
✓ **B. Summing over both subscripts i and j**
C. Multiplying two separate sums
D. Taking the average over i and j

Q30. For a demand function Qd = 50 − 2P and P = 10, the value of Qd is:

A. 50
✓ **B. 30**
C. 20
D. 70

Lecture 3: Quadratic, Exponential, Logarithmic Functions & Derivatives

Q31. A quadratic function has the general form:

A. Y = a + bX
✓ **B. $Y = aX^2 + bX + c$ where $a \neq 0$**
C. Y = a■
D. lnY = β0 + β1X

Q32. When the coefficient 'a' in the quadratic function $Y = aX^2 + bX + c$ is NEGATIVE, the graph:

A. Opens upward (U-shaped)
✓ **B. Opens downward (inverted U-shape)**
C. Is a straight horizontal line
D. Has no maximum or minimum

Q33. The Kuznets Curve describes the relationship between:

A. Trade and GDP
✓ **B. Income per capita and income inequality**
C. Population and fertility
D. Interest rates and investment

Q34. The Calmfors-Driffill hypothesis describes an inverted-U relationship between:

A. Inflation and unemployment
✓ **B. Degree of collective bargaining and level of unemployment**
C. GDP and exports
D. Capital and labor

Q35. Both Marginal Cost (MC) and Average Cost (AC) curves have which shape?

A. Inverted U-shape
B. Linear upward slope
✓ **C. U-shape**
D. Horizontal

Q36. The exponential function Y = a■ requires that:

A. a = 0 and x > 0
✓ **B. a > 0 and a ≠ 1**
C. a < 0
D. x = constant

Q37. Euler's constant 'e' used as the base of natural logarithm approximately equals:

A. 3.14159
✓ **B. 2.71828**
C. 1.41421
D. 1.61803

Q38. If Y = e■, then the natural logarithm of Y equals:

A. eY
✓ **B. x**
C. 1/x
D. Y/e

Q39. The concept of derivative represents:

A. The average value of a function
✓ **B. The instantaneous rate of change of the dependent variable**
C. The sum of all values of the function
D. The integral of a function

Q40. Using the Power Rule, if $y = 10x^3$, then dy/dx equals:

A. $10x^2$
✓ **B. $30x^2$**
C. $3x^2$
D. $30x^3$

Q41. The Constant Function Rule states that if y = k (a constant), then dy/dx equals:

A. k
B. 1
✓ **C. 0**
D. k^2

Q42. The Product Rule states that if y = f(x)·g(x), then dy/dx equals:

A. f'(x) + g'(x)
✓ **B. g(x)·f'(x) + f(x)·g'(x)**
C. f'(x)·g'(x)

D. f(x)/g'(x)

Q43. If y = f(x)/g(x), the Quotient Rule gives dy/dx as:

✓ **A. [g(x)·f'(x) – f(x)·g'(x)] / [g(x)]²**
B. [f(x)·g'(x) + g(x)·f'(x)] / [g(x)]²
C. f'(x)/g'(x)
D. g'(x)/f(x)

Q44. The Chain Rule is applied when:

A. A function involves only one variable
✓ **B. A function involves different variables (composition of functions)**
C. Both derivatives are zero
D. The function is linear

Q45. The marginal cost (MC) is mathematically defined as:

A. TC / Q
✓ **B. d(TC)/dQ**
C. TC × Q
D. 1/TC

Lecture 4: Partial Derivatives & Simple Optimization

Q46. Partial derivatives measure the rate of change of a dependent variable with respect to one independent variable while:

A. All other variables also change
✓ **B. All other independent variables are held constant**
C. The dependent variable is fixed
D. The function is differentiated twice

Q47. The mathematical symbol used to denote partial derivatives is:

A. Δ
B. d
✓ **C. ∂**
D. ∇

Q48. For Z = f(x, y) = 2x² + 3y² + 5xy + 20, the partial derivative ∂Z/∂x equals:

✓ **A. 4x + 5y**
B. 6y + 5x
C. 4x + 5
D. 2x + 3y

Q49. For Z = f(x, y) = 2x² + 3y² + 5xy + 20, the partial derivative ∂Z/∂y equals:

A. 4x + 5y
✓ **B. 6y + 5x**
C. 4y + 5
D. 2y + 5x

Q50. Cross partial derivatives Zxy and Zyx are always:

A. Opposite in sign
✓ **B. Equal (Symmetry of second derivatives)**
C. Different values
D. Zero

Q51. A 'Second Order Direct Partial Derivative' means differentiating:

A. Twice with respect to different variables
✓ **B. Twice with respect to the same variable**
C. Once with respect to each variable
D. The cross partial again

Q52. For optimization of a function of ONE variable, the First Order Condition requires:

A. f'(x) > 0
B. f'(x) < 0
✓ **C. f'(x) = 0**
D. f''(x) = 0

Q53. For a function of one variable, the Second Order Condition for MINIMIZATION requires:

A. f''(x) = 0
B. f''(x) < 0
✓ **C. f''(x) > 0**
D. f'(x) > 0

Q54. For a function of one variable, the Second Order Condition for MAXIMIZATION requires:

A. f''(x) > 0
✓ **B. f''(x) < 0**
C. f''(x) = 0
D. f'(x) = 0

Q55. For the profit function $\pi = 100Q - 120 - 2Q^2$, the profit-maximizing output Q is:

A. 100
B. 50
✓ **C. 25**
D. 120

Q56. What is the maximum profit when $\pi = 100Q - 120 - 2Q^2$ and Q = 25?

A. 2500
✓ **B. 1130**
C. 880
D. 1500

Q57. For multivariate functions, how many first-order partial derivatives are there?

A. Always two
✓ **B. As many as the number of independent variables**
C. Only one
D. Always three

Q58. A function Qd = f(P, I, Po, O) is called a:

A. Single variable function
✓ **B. Multivariate function**
C. Constant function
D. Discrete function

Q59. If $y = 40x - 2x^2$, the maximum value of y (at x = 10) is:

A. 400
✓ **B. 200**
C. 180

D. 300

Q60. The second derivative of $y = 10x^3$ is:

A. $30x^2$
✓ **B. $60x$**
C. $10x$
D. $90x^2$

Lecture 5: Multivariate Optimization & Probability Review

Q61. For multivariate optimization, the First Order Condition for a Local Minimum is:

A. $Zx > 0$ AND $Zy > 0$
✓ **B. $Zx = 0$ AND $Zy = 0$**
C. $Zx < 0$ AND $Zy < 0$
D. $Zxx > 0$ AND $Zyy < 0$

Q62. The Second Order Condition for a Local MINIMUM (multivariate) requires:

A. $Zxx < 0$ AND $Zyy < 0$
✓ **B. $Zxx > 0$ AND $Zyy > 0$**
C. $Zxx = 0$ AND $Zyy = 0$
D. $Zxx > 0$ AND $Zyy < 0$

Q63. The Second Order Condition for a Local MAXIMUM (multivariate) requires:

A. $Zxx > 0$ AND $Zyy > 0$
✓ **B. $Zxx < 0$ AND $Zyy < 0$**
C. $Zxx = 0$ AND $Zyy = 0$
D. $Zxx > 0$ AND $Zyy = 0$

Q64. A Saddle Point occurs when:

A. Both second order derivatives are positive
B. Both second order derivatives are negative
✓ **C. Second order derivatives have different signs**
D. The first order condition is not met

Q65. The Hessian condition to rule out points of inflection requires:

A. $Zxx \cdot Zyy < Zxy^2$
B. $Zxx \cdot Zyy = Zxy^2$
✓ **C. $Zxx \cdot Zyy > Zxy^2$**
D. $Zxx = Zyy$

Q66. A Random Experiment is defined as:

A. An experiment with exactly one possible outcome
✓ **B. Any process of observation with more than one possible outcome where we are uncertain about which will materialise**
C. A planned scientific experiment with controlled variables
D. An economic model with random parameters

Q67. The Sample Space is:

A. A subset of possible outcomes
✓ **B. The set of ALL possible outcomes of an experiment**
C. The probability of a specific event
D. The mean of a random variable

Q68. Mutually Exclusive Events means:

A. Both events occur simultaneously
✓ **B. Occurrence of one event prevents occurrence of the other at the same time**
C. Events are equally likely
D. Events exhaust all possible outcomes

Q69. The Classical Definition of Probability is:

A. Number of unfavourable outcomes / Total outcomes
✓ **B. Number of favourable outcomes / Total outcomes**
C. Number of trials / Number of events
D. 1 – P(not A)

Q70. A Continuous Random Variable:

A. Can take only whole number values
B. Has a finite number of possible values
✓ **C. Can take any value within an interval**
D. Is always normally distributed

Q71. The Normal Distribution has a curve that is:

A. Skewed to the right
B. Uniform (flat)
✓ **C. Bell-shaped with the highest point at the mean**
D. U-shaped

Q72. In a Normal Distribution, approximately what percentage of data falls within ±1 standard deviation of the mean?

A. 99.7%
B. 95%
✓ **C. 68%**
D. 50%

Q73. The Student's t-distribution compared to the Standard Normal distribution has:

A. Lighter tails (lower variance)
✓ **B. Heavier tails (higher variance)**
C. Exactly the same shape
D. Negative mean

Q74. The Chi-Square distribution is skewed and takes values:

A. From $-\infty$ to $+\infty$
✓ **B. From 0 to infinity**
C. From –1 to +1
D. From 0 to 1

Q75. The F-distribution is a:

A. Mean ratio distribution
✓ **B. Variance ratio distribution**
C. Standard deviation ratio distribution
D. Correlation ratio distribution

Lecture 6: Simple Regression Model & OLS

Q76. In regression analysis, the dependent variable Y is also known as the:

A. Regressor
B. Explanatory variable
✓ **C. Regressand**
D. Predictor variable

Q77. The independent variable X in regression is also called the:

A. Regressand
B. Response variable
✓ **C. Explanatory variable or Regressor**
D. Dependent variable

Q78. A scatter diagram is used BEFORE regression to:

A. Estimate the regression parameters
✓ **B. Observe the pattern and nature of relationship between variables**
C. Test the significance of coefficients
D. Calculate the coefficient of determination

Q79. The simple regression equation with an error term is written as:

A. $Y = \beta 0 + \beta 1X$
✓ **B. $Y = \beta 0 + \beta 1X + e$**
C. $Y = a + bX^2$
D. $Y = \beta 0 + \beta 1X - e$

Q80. The error term 'e' in the regression equation represents:

A. The explained variation in Y
✓ **B. The difference between actual Y and estimated Y**
C. The slope of the regression line
D. The value of the intercept

Q81. OLS stands for:

A. Optimal Linear Solution
✓ **B. Ordinary Least Squares**
C. Observed Linear Statistics
D. One-Level Summation

Q82. OLS minimises which of the following?

A. The sum of residuals
✓ **B. The sum of squared residuals**
C. The absolute value of residuals
D. The product of residuals

Q83. The first normal equation in OLS estimation is:

A. $\sum XY = a\sum X + b\sum X^2$
✓ **B. $\sum Y = na + b\sum X$**
C. $\sum Y^2 = na^2 + b^2\sum X^2$
D. $\sum e = a + b$

Q84. The second normal equation in OLS estimation is:

A. $\sum Y = na + b\sum X$
✓ **B. $\sum XY = a\sum X + b\sum X^2$**
C. $\sum X = na + b\sum Y$
D. $\sum e^2 = a\sum X + b\sum Y$

Q85. The formula for the slope coefficient 'b' in OLS is:

✓ **A. $b = (n\sum XY - \sum X \sum Y) / (n\sum X^2 - (\sum X)^2)$**

B. $b = \sum XY / \sum X^2$

C. $b = \sum Y / n - a$

D. $b = n\sum X / \sum Y$

Q86. The intercept 'a' can be found using:

A. a = ■ + bX■

✓ **B. a = ■ – bX■**

C. a = b × X■

D. $a = \sum Y/n + b$

Q87. Under the Gauss-Markov assumptions, OLS estimators are known as:

A. BLED

✓ **B. BLUE (Best Linear Unbiased Estimators)**

C. BLEF

D. BALE

Q88. Which of the following is a Gauss-Markov assumption?

A. Errors are heteroskedastic

✓ **B. Errors have zero conditional mean**

C. X values are all the same

D. Sample size is greater than 1000

Q89. Homoskedasticity means:

A. Errors have different variances for different observations

✓ **B. Errors have the same variance for all values of the explanatory variable**

C. Errors are correlated with each other

D. Errors are always positive

Q90. A model is 'linear' in the context of regression when:

A. Variables can only have power equal to one

✓ **B. The model is linear in parameters (slope coefficients)**

C. The graph is always a straight line in original form

D. The intercept must equal zero

Lecture 7: Estimation, Testing & R^2 in Regression

Q91. The intercept 'a' ($\beta 0$) in a regression equation typically represents:

A. The slope of the regression line

✓ **B. The value of Y when X is zero**

C. The coefficient of determination

D. The standard error of estimate

Q92. The slope coefficient 'b' ($\beta 1$) represents:

A. The value of Y when X = 0

✓ **B. The average unit change in Y per one-unit change in X**

C. The total variation in Y

D. The unexplained variation in Y

Q93. The Standard Error of Estimate (σ■e) is defined as:

A. $\sqrt{(\sum e^2 / n)}$

✓ **B. $\sqrt{(\sum e^2 / (N - k))}$**

C. $\sum e / n$

D. $\sum e^2 / (N - k)^2$

Q94. In the formula for Standard Error of Estimate, N – k represents:

A. Sample size minus one

✓ **B. Degrees of freedom**

C. Number of parameters

D. Coefficient of determination

Q95. The t-test for individual slope coefficient significance tests H0: b = 0 using the test statistic:

✓ **A. t = b / se(b)**

B. t = se(b) / b

C. t = b × se(b)

D. $t = b^2 / se(b)$

Q96. The region of rejection for the t-test (two-tailed) is:

A. $|t| < t\alpha/2, n–k$

✓ **B. $|t| > t\alpha/2, n–k$**

C. $t = t\alpha/2, n–k$

D. $t < 0$

Q97. 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 standard error of the regression

D. The number of observations

Q98. R^2 ranges between:

A. $–\infty$ and $+\infty$

B. –1 and +1

✓ **C. 0 and 1**

D. 0 and 100

Q99. The Goodness of Fit Test in regression uses the:

A. t-distribution

B. Normal distribution

✓ **C. F-distribution**

D. Chi-square distribution

Q100. For the F-test (Goodness of Fit), the region of rejection is:

A. $F < F\alpha, k–1, N–k$

✓ **B. $F > F\alpha, k–1, N–k$**

C. $F = F\alpha, k–1, N–k$

D. $F < 0$

Q101. If $R^2 = 0.934$, the F-statistic with n = 5 and k = 2 is approximately:

A. 15.3

✓ **B. 42.4**

C. 28.7

D. 9.34

Q102. The LINEST formula in Microsoft Excel returns regression statistics in which format?

A. A single cell
✓ **B. An array requiring Ctrl + Shift + Enter**
C. A chart
D. A separate worksheet automatically

Q103. The TREND formula in Excel computes:

A. The regression coefficients a and b
✓ **B. Estimated (trend) values of Y without calculating regression coefficients first**
C. The F-statistic
D. The coefficient of determination

Q104. In the Excel LINEST output, the first row contains:

A. Standard errors of coefficients
✓ **B. Coefficients b (slope) and a (intercept)**
C. R^2 and standard error of estimate
D. F-statistic and degrees of freedom

Q105. An 'Unbiased Estimator' means:

A. The estimator always gives the exact true value
✓ **B. The average value of the estimator in repeated samples equals the true population parameter**
C. The estimator has zero variance
D. The estimator is always positive

Lecture 8: Excel Applications in Regression

Q106. The Excel formula to find the slope coefficient 'b' is:

A. =INTERCEPT(Y range, X range)
✓ **B. =SLOPE(Y range, X range)**
C. =TREND(Y range, X range)
D. =LINEST(Y range)

Q107. The Excel formula to find the intercept 'a' is:

A. =SLOPE(Y range, X range)
✓ **B. =INTERCEPT(Y range, X range)**
C. =CORREL(Y range, X range)
D. =FORECAST(Y range, X range)

Q108. To read the critical t-value in Excel for a two-tailed test with α = 0.05 and 3 degrees of freedom, you type:

A. =T.INV(0.05, 3)
✓ **B. =T.INV.2T(0.05, 3)**
C. =T.DIST(0.05, 3)
D. =T.TEST(0.05, 3)

Q109. To read the critical F-value in Excel for F0.05, 1, 3 you type:

A. =F.INV(0.05, 1, 3)
✓ **B. =F.INV.RT(0.05, 1, 3)**
C. =F.TEST(0.05, 1, 3)
D. =F.DIST(0.05, 1, 3)

Q110. A linear transformation of data (like multiplying by a constant) does NOT change the:

A. Mean of the variable
B. Variance of the variable
✓ **C. Correlation between variables**
D. Scale of the variable

Q111. Non-linear transformations of data (like taking logs) DO change the:

A. Number of observations
✓ **B. Correlation between variables**
C. Names of variables
D. Source of the data

Q112. For the log-log model lnY = a + b lnX, the interpretation of 'b' is:

A. A one-unit change in X causes b units change in Y
✓ **B. A one-percent change in X causes b-percent change in Y (elasticity)**
C. A one-unit change in X causes 100b percent change in Y
D. b is the intercept of the original function

Q113. For the log-level model lnY = a + bX, when X changes by 1 unit, Y changes by approximately:

A. b units
B. b percent
✓ **C. b × 100 percent**
D. b/X percent

Q114. For the level-log model Y = a + b lnX, when X changes by 1 percent, Y changes by:

A. b units
✓ **B. b/100 units**
C. b × 100 units
D. b × X units

Q115. An Outlier in regression is:

A. A variable with a high coefficient
✓ **B. A data point that diverges significantly from the general pattern and may affect the slope**
C. An observation with zero residual
D. An independent variable with low correlation

Q116. An 'Influential Point' is an outlier that:

A. Has no effect on the regression line
✓ **B. Changes the slope or R^2 significantly when removed**
C. Is always caused by data entry errors
D. Is always at the highest X value

Q117. The standard error of regression σ■e gives a measure of:

A. The slope steepness
✓ **B. The accuracy of predictions**
C. The number of outliers
D. The intercept value

Q118. For the Cobb-Douglas function lnY = lnA + α lnL + β lnK, what does α represent?

A. The output elasticity with respect to Capital
✓ **B. The output elasticity with respect to Labour**
C. The total factor productivity

D. The intercept of the linear form

Q119. Change of unit of measurement in regression changes:

A. The correlation between X and Y
B. The sign of the slope coefficient
✓ C. The numerical values of coefficients but not their significance
D. The R-squared value fundamentally

Q120. In a scatter diagram, if Y increases with X but not perfectly, this is best captured by:

✓ A. Y = a + bX (the OLS regression equation with error term)
B. Y = aX (no intercept)
C. Y = a constant
D. lnY = a

Lecture 9: Multiple Regression

Q121. Multiple Regression differs from Simple Regression in that it considers:

A. Only one explanatory variable
✓ B. Two or more explanatory variables simultaneously
C. Only time series data
D. Only qualitative variables

Q122. The Multiple Regression equation with two independent variables is written as:

A. $Yi = \beta0 + \beta1X1i + ei$
✓ B. $Yi = \beta0 + \beta1X1i + \beta2X2i + ei$
C. $Yi = \beta1X1i + \beta2X2i$
D. $Yi = \beta0 + \beta1(X1 + X2) + ei$

Q123. Multiple Regression with two independent variables fits the data in:

A. One-dimensional space
B. Two-dimensional space
✓ C. Three-dimensional space
D. Four-dimensional space

Q124. For Multiple Regression with K independent variables, OLS minimization produces how many normal equations?

A. K
✓ B. K + 1
C. K – 1
D. 2K

Q125. In the deviation form for multiple regression, lowercase x1 represents:

A. The original value of X1
✓ B. X1i – X■1 (deviation from mean)
C. lnX1
D. X1 squared

Q126. The expression for $\beta1$ in multiple regression (deviation form) involves:

A. Only $\sum x1y$
✓ B. $\sum x1y$, $\sum x2^22$, $\sum x2y$, $\sum x1x2$, $\sum x1^2$
C. Only $\sum XY$ and $\sum X^2$
D. Only the correlation between X1 and Y

Q127. After adding a new independent variable to a regression, the R^2 typically:

A. Decreases
B. Stays the same
✓ **C. Increases**
D. Becomes zero

Q128. The degree of freedom in multiple regression with N observations and K parameters is:

A. N + K
B. N × K
✓ **C. N – K**
D. K – N

Q129. In multiple regression, a variable X1 is considered significant if the t-test shows:

A. t = 0
✓ **B. $|t| > t\alpha/2, n–k$**
C. $|t| < t\alpha/2, n–k$
D. t > 0 always

Q130. The F-test in multiple regression tests:

A. The significance of individual slope coefficients
✓ **B. Whether all slope coefficients are simultaneously zero**
C. The presence of outliers
D. The degree of multicollinearity

Q131. For multiple regression with LINEST in Excel, you need to select a cell array with:

A. 1 column and 5 rows regardless of variables
✓ **B. (k+1) columns and 5 rows, where k is number of independent variables**
C. Always 2 columns
D. n rows and 3 columns

Q132. The conversion formula $\sum x1y = \sum X1Y – (\sum X1 \sum Y)/n$ converts summations from:

A. Deviation form to normal form
✓ **B. Normal form to deviation form**
C. Log form to level form
D. Cross-section to time series

Q133. The formula $\beta 0 = \blacksquare – \beta 1X\blacksquare 1 – \beta 2X\blacksquare 2$ gives the value of:

A. The slope for X1
✓ **B. The intercept in multiple regression**
C. The R-squared value
D. The F-statistic

Q134. In the multiple regression LINEST output, the order of coefficients displayed is:

A. $\beta 0, \beta 1, \beta 2$ from left to right
✓ **B. $\beta 2, \beta 1, \beta 0$ from left to right (reverse order)**
C. $\beta 1, \beta 2, \beta 0$
D. Random order

Q135. If $\beta 2$ is not significant in multiple regression, it means:

A. X2 has a strong impact on Y
✓ **B. X2 does not seem to have a significant independent impact on Y**
C. The model has zero R^2

D. The regression should be discarded entirely

Lecture 10: K-Variable Regression & Analysis ToolPak

Q136. The Analysis ToolPak in Excel restricts the number of independent variables to a maximum of:

A. 50
✓ **B. 16**
C. 10
D. 100

Q137. To install the Analysis ToolPak in Excel 2010, you go to:

A. Insert → Analysis ToolPak
✓ **B. File → Options → Add-Ins → Manage Excel Add-Ins → GO**
C. View → Data Analysis
D. Home → Add-Ins

Q138. In the Analysis ToolPak regression output, 'P-value' represents:

A. The probability that the null hypothesis is true
✓ **B. The level of significance at which the t-statistic becomes significant**
C. The coefficient of determination
D. The F-statistic

Q139. For the K-variable linear regression, the number of independent variables is denoted as:

A. n
✓ **B. K (where K < M and M is finite)**
C. N
D. e

Q140. The Adjusted R^2 formula adjusts R^2 by:

A. Adding the number of variables
✓ **B. Penalising for adding more independent variables by accounting for degrees of freedom**
C. Multiplying by sample size
D. Dividing by the number of observations

Q141. Adjusted $R^2 = R^2 - (1 - R^2) \times [(k - 1)/(N - k)]$ is used because:

✓ **A. R^2 always overestimates when more variables are added**
B. R^2 is always negative
C. We want to increase the R^2 artificially
D. The standard R^2 is undefined for multiple regression

Q142. Multiple Correlation is equal to:

A. R^2
✓ **B. $\sqrt{R^2}$**
C. $(R^2)^2$
D. $1 - R^2$

Q143. A key LIMITATION of Excel's Data Analysis ToolPak is:

A. It cannot compute R^2
✓ **B. It does not provide robust standard errors or alternative model options**
C. It requires STATA to run
D. It only works for simple regression

Q144. In the Data Analysis regression window, you must select all X-variable columns that are:

A. In separate worksheets
✓ **B. In adjoining (adjacent) columns**
C. In non-adjacent columns
D. Sorted in ascending order

Q145. The Analysis ToolPak regression output shows the ANOVA table, which displays:

A. Individual variable significance
✓ **B. Sums of Squares (SS) and Mean Squares (MS) and the F-statistic**
C. Only the intercept
D. The scatter diagram

Q146. As the number of independent variables increases in a regression model:

A. Degrees of freedom increase
✓ **B. Degrees of freedom decrease**
C. R^2 must decrease
D. The F-statistic is always larger

Q147. Real-world regression on Pakistan GDP data mentioned in the lecture used data from:

A. State Bank of Pakistan only
✓ **B. World Development Indicators (World Bank)**
C. IMF database
D. Pakistan Bureau of Statistics

Q148. The P-value for a coefficient equals approximately zero means:

A. The coefficient is not significant
✓ **B. The coefficient is significant at very low levels of significance (e.g., below 1%)**
C. The R^2 is zero
D. The F-test fails

Q149. To run regression using Analysis ToolPak, click on the DATA ribbon and then select:

A. Solver
✓ **B. Data Analysis**
C. What-If Analysis
D. Forecast Sheet

Q150. After entering LINEST formula, you must press:

A. Enter only
✓ **B. Ctrl + Shift + Enter**
C. Alt + F4
D. Ctrl + C

Lecture 11: Transformation & Functional Forms in Regression

Q151. 'Linear in the linear regression' means linear in:

A. Variables only
✓ **B. Parameters (coefficients), not necessarily in variables**
C. Both variables and parameters
D. The error term

Q152. Which functional form is appropriate when the dependent variable is expected to increase at a DECREASING rate with respect to X?

A. Linear: Y = a + bX

✓ **B. Logarithmic: Y = a + b lnX**

C. Quadratic with positive a: Y = a + bX²

D. None of the above

Q153. U-shaped curves are used to model which economic relationship?

A. Total Revenue curve

✓ **B. Marginal Cost and Average Cost curves**

C. Linear demand curves

D. Exponential supply curves

Q154. A Total Cost curve with an 'S-shape' can be represented by a:

A. Quadratic function

✓ **B. Cubic function**

C. Double-log function

D. Linear function

Q155. When Y is a function of X and we add an interaction term (X■ × X■), it captures:

A. The simple additive effect of both X■ and X■

✓ **B. The effect of X■ on Y that depends on the value of X■**

C. The correlation between X■ and X■

D. The outlier effect

Q156. In the marks-IQ-hardwork example M = a + bI + cH + d(I × H) + e, the coefficient 'd' represents:

A. The direct effect of IQ on marks

✓ **B. The additional effect of having both IQ and hard work combined (interaction)**

C. The average hard work level

D. The residual effect

Q157. A Dynamic Model with a lagged independent variable is written as:

A. Yt = a + bXt

✓ **B. Yt = a + bXt + cXt−1 + e (X observed in previous period)**

C. Yt = a + bXt²

D. Yt = lnXt + e

Q158. The agricultural supply model Qs = a + bPt−1 + e shows that supply depends on:

A. The current price

✓ **B. The price of the PREVIOUS time period**

C. The price two periods ahead

D. The average price

Q159. The scatter plot showing a LOGARITHMIC shape suggests using:

A. Y = a + bX

✓ **B. Y = a + b lnX**

C. Y = aX²

D. Y = a × e■

Q160. For the polynomial model Y = β0 + β1X + β2X², if β2 > 0, the curve is:

A. Inverted-U shape

✓ **B. U-shaped (minimum exists)**

C. Linear

D. Exponential

Q161. Diminishing returns to a factor of production can be modeled using:

A. A linear form
✓ B. A logarithmic or polynomial (cubic) form
C. A constant function
D. A step function

Q162. In the model lnY = β0 + β1X (log-level/semi-log), the marginal effect of X on Y is:

A. β1
✓ B. β1Y
C. β1/Y
D. β1/X

Q163. For the linear-in-variable model Y = a + bX, the elasticity ε equals:

A. b
✓ B. b × (X/Y)
C. b × Y/X
D. b^2

Q164. The model Y = AKβ (Cobb-Douglas in capital only) after taking logs becomes:

A. Y = A + βK
✓ B. lnY = lnA + β lnK
C. Y = A × βlnK
D. lnY = A × lnK

Q165. Choosing between different functional forms should primarily be guided by:

A. The size of R^2
✓ B. Theory and scatter diagram indicating the nature of the relationship
C. The number of parameters
D. The personal preference of the researcher

Lecture 12: Standardised Variables & Regression Strategy

Q166. A standardised variable X* is computed as:

✓ A. (Xi – X■) / SX
B. (Xi × SX) / X■
C. Xi / SX
D. Xi – SX

Q167. After standardisation, the mean of the standardised variable is:

A. 1
✓ B. 0
C. SX
D. X■

Q168. After standardisation, the variance of the standardised variable is:

A. 0
B. SX
✓ C. 1
D. n

Q169. When a regression is run on standardised variables, the intercept 'a' is always:

A. Equal to X■

✓ **B. Zero**

C. Equal to β1

D. Undefined

Q170. Beta coefficients (standardised coefficients) allow comparison of:

A. Variables in the same units

✓ **B. The relative importance of predictors regardless of their units**

C. The raw coefficients across models

D. The R^2 of different models

Q171. The interpretation of a Beta coefficient of 0.8 for X1 is:

A. A one-unit change in X1 changes Y by 0.8 units

✓ **B. A one standard deviation increase in X1 is associated with a 0.8 standard deviation increase in Y**

C. X1 accounts for 80% of R^2

D. X1 is significant at the 80% level

Q172. A disadvantage of standardised coefficients is:

A. They are harder to compute

✓ **B. The meaning of a standard deviation varies between non-normal distributions**

C. They cannot be used for hypothesis testing

D. They always equal zero

Q173. Steps BEFORE regression analysis include all of the following EXCEPT:

A. Theory

B. Graphs

C. Summary Statistics

✓ **D. Publishing the paper**

Q174. Summary statistics in Excel Analysis ToolPak are obtained by selecting:

A. Regression

✓ **B. Descriptive Statistics**

C. Correlation

D. Histogram

Q175. The correlation matrix in Excel Analysis ToolPak is useful because:

A. It estimates regression coefficients

✓ **B. It helps identify relationships between variables, including potential multicollinearity**

C. It computes the F-statistic

D. It removes outliers

Q176. A strategy for Keeping or Dropping Variables in regression suggests:

A. Always drop variables with unexpected signs

✓ **B. Keep significant predictors and insignificant predictors with expected signs; drop insignificant ones with unexpected signs**

C. Keep all variables regardless of significance

D. Drop all variables with p-values above 0.01

Q177. To present regression results in a research paper, you typically report:

A. Only the R^2 value

✓ **B. Coefficients, standard errors, significance indicators, R^2, observations, and the dependent variable**

C. Only the slope coefficients

D. Only the F-statistic

Q178. It is common practice in research papers to present:

A. Only one regression model

✓ **B. Multiple regression models (e.g., Model 1, Model 2) to show robustness**

C. The data in full

D. The scatter diagrams without regression

Q179. The RSQ function in Excel computes:

✓ **A. R-squared**

B. Standard error

C. Slope

D. Intercept

Q180. The LINEST function can handle up to _____ independent variables:

A. 16

✓ **B. 50**

C. 10

D. 100

Lecture 13: Qualitative Variables & Dummy Variables

Q181. A Dummy Variable is also known as all of the following EXCEPT:

A. Binary variable

B. Indicator variable

C. Qualitative independent variable

✓ **D. Continuous variable**

Q182. A binary dummy variable takes values:

✓ **A. 0 and 1 only**

B. 1 and 2 only

C. Any positive number

D. –1 and 0

Q183. If G = 1 for female and G = 0 for male in the regression $W = \beta 0 + \beta 1E + \beta 2X + \beta 3G + e$, then $\beta 3$ represents:

A. The effect of education on wages

✓ **B. The average wage difference between females and males, holding E and X constant**

C. The intercept for male wages

D. The slope of experience on wages

Q184. A negative $\beta 3$ for the female dummy variable means:

A. Females earn more than males

✓ **B. Females earn, on average, $|\beta 3|$ less than males (ceteris paribus)**

C. Gender has no effect on wages

D. The model is incorrectly specified

Q185. For a qualitative variable with THREE categories (e.g., Race: White, Black, Asian), how many dummy variables should be included in regression?

A. 3 (one for each category)
✓ **B. 2 (one less than the number of categories)**
C. 1
D. 4

Q186. The category for which all dummy variables are set to zero is called the:

A. Treatment variable
✓ **B. Reference variable (or base category)**
C. Dependent variable
D. Interaction variable

Q187. If we include ALL categories as dummy variables in regression (e.g., all three for Race), we encounter:

A. Heteroskedasticity
✓ **B. The Dummy Variable Trap (perfect multicollinearity)**
C. Autocorrelation
D. Missing data

Q188. An interaction term between a dummy variable G and continuous variable E (G × E) captures:

A. The sum of effects of G and E
✓ **B. The difference in the slope of E for different values of G**
C. The average value of wages
D. The correlation between G and E

Q189. For the model W = β0 + β1E + β2X + β3G + β4(G×E) + e, for WOMEN (G=1) the effective intercept is:

A. β0
✓ **B. β0 + β3**
C. β0 + β4
D. β3 + β4

Q190. For the model W = β0 + β1E + β2X + β3G + β4(G×E) + e, for WOMEN (G=1) the effective slope on E is:

A. β1
✓ **B. β1 + β4**
C. β1 × β4
D. β4

Q191. The result W = 77.76 + 8.29E + 28.52X – 26.6G means that with identical education and experience, females earn:

A. 26.6 more than males
✓ **B. 26.6 less than males**
C. 77.76 less than males
D. The same as males

Q192. Dummy variables are called 'Polytomous' when:

A. They take values 0 and 1
✓ **B. The qualitative variable has MORE than two categories**
C. They are used for interaction only
D. They are continuous variables

Q193. If D1 = 1 for Black and D2 = 1 for Asian (White is reference), then for a Black Male with G = 0:

✓ **A. $W = \beta0 + \beta1E + \beta2X + \beta3(0) + \beta4(1) + \beta5(0) + e = \beta0 + \beta4 + \beta1E + \beta2X + e$**

B. $W = \beta0 + \beta3 + \beta1E + \beta2X + e$

C. $W = \beta0 + \beta5 + \beta1E + \beta2X + e$

D. $W = \beta0 + \beta1E + e$

Q194. Interaction between race and gender dummies is used when:

A. Race and gender are correlated

✓ **B. The effect of race on wages differs between males and females**

C. There are missing values

D. The model has only one dummy variable

Q195. The 'Dummy Variable Trap' refers to the problem of:

A. Having too few dummy variables

✓ **B. Including perfectly collinear dummy variables (all categories) causing estimation failure**

C. Using dummy variables for continuous data

D. Having negative dummy variable coefficients

Lecture 14: Logarithms in Regression & Missing Data

Q196. Using the natural log of a dependent variable is recommended when:

A. The variable is already normally distributed

✓ **B. The variable is positively skewed**

C. The variable is negatively skewed

D. The variable is binary

Q197. MCAR stands for:

A. Missing Cannot be Analyzed for Regression

✓ **B. Missing Completely at Random**

C. Multiple Correlation Adjustment in Regression

D. Missing Categorical and Random

Q198. In MCAR, the probability of a value being missing:

A. Depends on both X and Y

B. Depends only on X

✓ **C. Depends on neither X nor Y**

D. Depends only on Y

Q199. MAR (Missing at Random) means:

A. Probability of missingness depends on both X and Y

✓ **B. Probability of missingness depends on X but NOT on Y**

C. Probability of missingness does not depend on anything

D. Probability of missingness depends only on Y

Q200. MNAR (Missing Not at Random) is the most problematic because:

A. It is easy to handle with listwise deletion

✓ **B. The probability of missing depends on the unobserved value itself, making estimates biased**

C. It reduces the sample to zero

D. It only occurs in time series

Q201. Listwise Deletion (complete case analysis) means:

A. Deleting only certain variables
✓ **B. Deleting any row that has at least one missing value**
C. Substituting mean values for missing data
D. Running separate regressions for each missing pattern

Q202. The disadvantage of Listwise Deletion is:

A. It is too complex to implement
✓ **B. It reduces statistical power by using fewer observations**
C. It biases the slope coefficients upward
D. It requires the data to be MNAR

Q203. Pairwise Deletion (available case analysis) uses:

✓ **A. All available observations for each specific analysis**
B. Only complete rows for all analyses
C. Only the first and last observations
D. Only median values for missing cases

Q204. A disadvantage of Pairwise Deletion is:

A. It wastes all available data
✓ **B. The sample may differ across different analyses within the same study**
C. It requires MNAR assumption
D. It is computationally impossible

Q205. Substituting the MEAN for missing values in Single Imputation leads to:

A. Increased variability
✓ **B. Weakened covariance and reduced variability**
C. Perfect estimation of missing values
D. No change in statistical results

Q206. Multiple Imputation is preferred over Single Imputation because:

A. It is simpler to implement
✓ **B. It generates multiple datasets allowing better estimation of variability and reducing bias**
C. It always gives exactly the true missing value
D. It requires only one regression model

Q207. Maximum Likelihood imputation identifies parameters that:

A. Minimise the sum of squared residuals
✓ **B. Produce the maximum log-likelihood given the observed data**
C. Equal the sample mean
D. Minimise the number of missing observations

Q208. Attrition in panel data (like death, dropout) is a cause of missing data due to:

A. Measurement errors
✓ **B. Natural process attrition**
C. Wrong model specification
D. Multicollinearity

Q209. Using logs of wages in regression increases R^2 because:

A. Logs artificially inflate R^2
✓ **B. Log-transformation of skewed variables better meets regression assumptions**
C. Logs convert all values to 1
D. Logs eliminate the intercept

Q210. In the regression of ln(Wage) on ln(tenure), the coefficient on ln(tenure) equals the:

A. Marginal effect
✓ **B. Wage elasticity with respect to tenure**
C. Standard error
D. Growth rate of wages

Lecture 15: Multicollinearity

Q211. Perfect Multicollinearity exists when:

A. Error terms are correlated
✓ **B. There is an EXACT linear relationship between explanatory variables**
C. The dependent variable is correlated with residuals
D. The sample size is too small

Q212. Near or Imperfect Multicollinearity exists when:

A. There is an exact linear relationship between regressors
✓ **B. There is a NEAR or high (but not perfect) linear relationship between explanatory variables**
C. R^2 equals exactly 1
D. The dependent variable is binary

Q213. Under perfect multicollinearity (e.g., X1 = 5X2), what does the estimation software typically do?

A. Produces very large coefficients
✓ **B. Drops one of the collinear variables and shows a zero coefficient for it**
C. Produces negative R^2
D. Automatically corrects the problem

Q214. Which of the following is a SOURCE of Multicollinearity?

A. Large sample size
✓ **B. Common time series trend among explanatory variables**
C. Use of dummy variables
D. High R^2 of the model

Q215. A practical consequence of Multicollinearity includes:

A. Low R^2 but high t-statistics
✓ **B. High R^2 but low t-statistics (insignificant coefficients)**
C. Small standard errors
D. Biased R^2

Q216. Under Multicollinearity, OLS estimators are:

A. Biased
✓ **B. Still BLUE (unbiased) but have larger variances**
C. Efficient but biased
D. Inconsistent

Q217. VIF stands for:

A. Variable Inclusion Factor
✓ **B. Variance Inflation Factor**
C. Variable Interaction Function
D. Variance Independence Factor

Q218. Tolerance is defined as:

A. $1 / (1 – R^2)$
✓ **B. $1 – R^2$**
C. $R^2 / (1 – R^2)$
D. $1 / R^2$

Q219. A VIF value ≥ 10 indicates:

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

Q220. To detect Multicollinearity, we can look at the Correlation Matrix. High pairwise correlation between explanatory variables (e.g., > 0.8) suggests:

A. Homoskedasticity
✓ **B. Possible multicollinearity**
C. Good model specification
D. No issues with the data

Q221. Auxiliary regressions for detecting Multicollinearity involve:

A. Regressing Y on each X separately
✓ **B. Regressing each Xi on all other Xj variables and examining R^2**
C. Regressing residuals on Y
D. Regressing Y on squared terms

Q222. Which symptom is a tell-tale sign of multicollinearity?

A. Low R^2 with significant t-statistics
✓ **B. High R^2 with very low t-statistics (coefficients not significant)**
C. Negative R^2
D. Very small standard errors

Q223. The over-determined model (large number of explanatory variables with very low degrees of freedom) is a source of:

A. Heteroskedasticity
✓ **B. Multicollinearity**
C. Autocorrelation
D. Omitted variable bias

Q224. Multicollinearity is fundamentally described as a:

A. Model specification error
✓ **B. Data deficiency problem**
C. Measurement error problem
D. Sampling bias problem

Q225. Which of the following is NOT listed as a consequence of Multicollinearity?

A. Wide confidence intervals
✓ **B. Estimators become biased**
C. Wrong signs of coefficients
D. High sensitivity to data changes

Lecture 16: Multicollinearity: Remedial Measures

Q226. Which remedial measure for Multicollinearity involves not doing anything?

A. Dropping a variable
B. Transforming variables
✓ **C. Doing nothing (especially if MC is mild or purpose is forecasting)**
D. Ridge regression

Q227. Dropping a variable to remedy Multicollinearity risks:

A. Heteroskedasticity
✓ **B. Omitted variable bias / wrong model specification**
C. Autocorrelation
D. Perfect R^2

Q228. Combining exports and imports into a single 'openness' variable is an example of which remedial measure?

A. Dropping a variable
✓ **B. Transforming variables to reduce multicollinearity**
C. Using ridge regression
D. Increasing sample size

Q229. The first difference form ($Y_t - Y_{t-1}$) for time series is used to remedy multicollinearity but has the cost of:

A. Increasing R^2
✓ **B. Losing one observation**
C. Making all coefficients negative
D. Making R^2 undefined

Q230. In the Multicollinearity example with 10 observations and prices of two substitutes, the VIF for X1 was approximately:

A. 5
B. 10
✓ **C. 500**
D. 2

Q231. After increasing the sample to 50 observations in the multicollinearity example, the VIF for prices of substitutes:

A. Disappeared (became 1)
✓ **B. Decreased from ~500 to ~13.7 but MC still existed**
C. Increased further
D. Became undefined

Q232. When one of two multicollinear price-of-substitute variables is dropped from the model:

A. R^2 increases dramatically
✓ **B. R^2 remains similar but VIF decreases (no MC) and significance improves**
C. The remaining coefficient becomes insignificant
D. The intercept becomes zero

Q233. Using Panel data is listed as a remedial measure for multicollinearity because:

A. Panel data has no explanatory variables
✓ **B. Combining cross-sectional and time series variation can reduce MC**
C. Panel data always reduces the number of variables
D. Panel data eliminates the intercept

Q234. Ridge regression is mentioned as a remedial option for multicollinearity that belongs to which category?

A. Dropping variables
✓ **B. Advanced estimation techniques beyond standard OLS**
C. Data transformation
D. Increasing sample size

Q235. In the GDP example with Pakistan data, after dropping imports, the situation improved because:

A. All variables became perfectly collinear
✓ **B. VIF values decreased and coefficients became more significant with correct signs**
C. R^2 decreased significantly
D. The intercept became zero

Q236. Which approach for remedying multicollinearity is appropriate when theory requires a specific form?

A. Drop the theoretically important variable
✓ **B. Do nothing or seek more data as omitting it would cause misspecification**
C. Replace it with its square
D. Use only the interaction term instead

Q237. Factor analysis is listed as a technique to:

A. Test for heteroskedasticity
✓ **B. Remedy multicollinearity**
C. Detect autocorrelation
D. Impute missing data

Q238. In the GDP-Pakistan multicollinearity example, combining exports and imports as 'Openness' is classified as:

A. Ridge regression
✓ **B. Variable transformation**
C. Listwise deletion
D. Dummy variable approach

Q239. When MC arises due to common time series trends, the first-difference transformation is useful because:

A. It converts time series to cross-sectional data
✓ **B. It removes common trends making variables less collinear**
C. It doubles the sample size
D. It eliminates the need for regression

Q240. A model is said to be 'over-determined' when:

A. There are more observations than variables
✓ **B. There are more explanatory variables than observations, giving very low degrees of freedom**
C. The intercept is zero
D. R^2 is exactly 1

Lecture 17: Heteroskedasticity: Causes & Detection

Q241. Homoskedasticity requires that:

A. $Var(e_i|X) = \sigma^2_i$ (varies across observations)

✓ **B. Var(ei|X) = σ² (constant for all observations)**

C. Errors are correlated

D. The mean of errors is non-zero

Q242. Heteroskedasticity means:

A. Errors have constant variance

✓ **B. Errors have non-constant (variable) variance across observations**

C. Errors are correlated across observations

D. The model has wrong functional form

Q243. 'Hetero' in Heteroskedasticity means:

A. Same

✓ **B. Different**

C. Random

D. Large

Q244. Which type of data is MORE LIKELY to exhibit Heteroskedasticity?

A. Time series data

✓ **B. Cross-sectional data**

C. Panel data with few time periods

D. Balanced panel data

Q245. Example of Heteroskedasticity: Expenditure on vacation = f(family income). Why is HSK likely?

A. Low-income families have more vacation options

✓ **B. High-income families have greater variability in vacation spending**

C. All families spend the same on vacations

D. Income is always exactly measured

Q246. Which of the following is NOT a cause of Heteroskedasticity listed in the lecture?

A. Variance of errors increasing with explanatory variable

B. Wrong model specification

✓ **C. Large sample size**

D. Measurement errors

Q247. Under Heteroskedasticity, OLS estimators are:

A. Biased and inconsistent

✓ **B. Unbiased and consistent but INEFFICIENT (not minimum variance)**

C. Best Linear Unbiased Estimators

D. Efficient but biased

Q248. A key consequence of Heteroskedasticity for inference is:

A. Slope coefficients become zero

✓ **B. Standard errors and confidence intervals are incorrect, leading to wrong t and F conclusions**

C. R^2 increases to 1

D. The intercept becomes unidentifiable

Q249. To DETECT Heteroskedasticity visually, we typically plot:

A. Y against time

✓ **B. Residuals (or squared residuals) against each independent variable**

C. Y against its mean

D. X1 against X2

Q250. Under Homoskedasticity, the scatter of residuals against X shows:

A. A systematic fan or funnel shape

✓ B. A random pattern with constant spread (no trend in variability)

C. A perfectly linear trend

D. All residuals equal to zero

Q251. Under Heteroskedasticity, the plot of residuals squared against X may show:

A. A random horizontal band

✓ B. A systematic pattern such as a funnel or increasing spread

C. No visible pattern

D. Residuals decreasing uniformly

Q252. Wrongly specified model (e.g., omitting a squared term) can cause:

A. Multicollinearity only

✓ B. Heteroskedasticity (and other problems)

C. Perfect fit

D. Zero residuals

Q253. Measurement errors in data can cause Heteroskedasticity because:

A. Perfect data is always collected

✓ B. Some respondents provide less accurate information systematically (e.g., high income groups underreporting)

C. Measurement errors reduce the sample size

D. Measurement errors cancel each other out

Q254. Under Heteroskedasticity, OLS gives more weight to observations with:

A. Smaller errors

✓ B. Larger errors (minimising the sum of squared errors overemphasises extreme values)

C. Median values

D. Zero residuals

Q255. Which form of data is mentioned as facing Heteroskedasticity due to seasonal components?

A. Pure cross-section data

✓ B. Time series data with seasonal patterns

C. Randomly sampled cross-sectional data

D. Perfectly balanced panel data

Lecture 18: Heteroskedasticity: Formal Tests

Q256. The Goldfeld-Quandt Test for Heteroskedasticity involves:

A. Running one regression on the full sample

✓ B. Splitting the data into high and low groups (omitting middle), running separate regressions, and comparing RSS

C. Regressing squared residuals on all regressors

D. Running auxiliary regressions for each X variable

Q257. In the Goldfeld-Quandt Test, the middle C observations are omitted because:

A. They are outliers

✓ B. To create a clearer distinction between high-variance and low-variance groups

C. They are always missing

D. They are perfectly collinear

Q258. The test statistic in the Goldfeld-Quandt Test is:

A. t = b / se(b)
✓ B. F = RSS_high / RSS_low
C. LM = N × R^2
D. χ^2 = N × R^2

Q259. In the Goldfeld-Quandt Test, the degrees of freedom for numerator and denominator are both:

A. N – K
✓ B. (N – C)/2 – K
C. N/2 – 1
D. C/2 – K

Q260. A drawback of the Goldfeld-Quandt Test is:

A. It is too sensitive to normality violations
✓ B. It cannot handle situations where SEVERAL variables jointly cause HSK
C. It cannot detect HSK at all
D. It requires a very large sample of over 1000

Q261. The Park LM Test for Heteroskedasticity involves regressing:

A. e^2 on Y
✓ B. ln e^2 on ln X1, ln X2, ..., ln XK
C. e on X1, X2, ..., XK
D. |e| on X1, X2, ..., XK

Q262. In the Park LM Test, the test statistic is:

A. F = RSS_high / RSS_low
✓ B. LM = N × R^2 from the auxiliary regression
C. t = b / se(b)
D. The Durbin-Watson statistic

Q263. The Park LM Test and Harvey-Godfrey Test are sensitive to:

A. Having too few variables
✓ B. Violations of the normality assumption of errors
C. Large sample sizes
D. Negative R^2 values

Q264. The Breusch-Pagan Test for Heteroskedasticity regresses:

A. ln e^2 on ln X1, ..., ln XK
✓ B. e^2 on X1, X2, ..., XK (the original explanatory variables)
C. |e| on √X1
D. Y on e^2

Q265. The White Test is considered the MOST POPULAR test for Heteroskedasticity because:

A. It is the oldest test
✓ B. It includes squared terms and cross-products, making it more general and robust to non-normality
C. It requires the fewest steps
D. It always detects HSK

Q266. In the White Test, the auxiliary regression includes:

A. Only the original X variables
✓ **B. Original X variables, their squares, and cross-products**
C. Only the squared residuals regressed on Y
D. Only the lagged values of X

Q267. The LM test statistic ($N \times R^2$) is compared against the:

A. t-distribution
B. F-distribution
✓ **C. Chi-square distribution**
D. Normal distribution

Q268. The Glesjer Test for Heteroskedasticity regresses:

A. Squared residuals on X
✓ **B. Absolute residuals |ei| on various transformations of X (like X, √X, 1/X)**
C. ln(residuals) on ln(X)
D. Y on absolute residuals

Q269. If LM > χ^2K−1 in a Heteroskedasticity test, we:

A. Accept the null hypothesis of Homoskedasticity
✓ **B. Reject the null hypothesis and conclude Heteroskedasticity exists**
C. Conclude perfect collinearity
D. Cannot draw any conclusion

Q270. Which of the following best summarises the consequences of Heteroskedasticity on OLS?

A. Coefficients are biased and t-tests are correct
✓ **B. Coefficients are unbiased but standard errors are wrong, making inference unreliable**
C. Coefficients are biased and R^2 becomes negative
D. Coefficients and standard errors are both correct

Lecture 1: Overview of Econometrics & Data Types

Q271. Which type of data would you use to study how the sales of the SAME 100 companies changed over 10 years?

A. Cross-sectional data
✓ **B. Panel (Longitudinal) data**
C. Pooled Cross-Sectional data
D. Simple time series data (not panel)

Lecture 11: Transformation & Functional Forms in Regression

Q272. The statement 'A one percent increase in X leads to a β1 percent increase in Y' is the interpretation of which functional form?

A. Linear: Y = a + bX
✓ **B. Log-Log (Double-log): lnY = a + b lnX**
C. Log-Level: lnY = a + bX
D. Level-Log: Y = a + b lnX

Lecture 7: Estimation, Testing & R^2 in Regression

Q273. In hypothesis testing, the probability of rejecting a true null hypothesis is known as:

A. Power of the test

B. Type II error

✓ **C. Level of significance (Type I error)**

D. P-value of the alternative hypothesis

Lecture 6: Simple Regression Model & OLS

Q274. The residual sum of squares (RSS) in OLS equals which of the following?

A. $\sum(Yi - ■)^2$

B. $\sum(■i - ■)^2$

✓ **C. $\sum(Yi - ■i)^2$**

D. $\sum(Yi - \beta0)^2$

Lecture 7: Estimation, Testing & R^2 in Regression

Q275. What does $R^2 = 1$ mean in regression?

A. No explanatory power

B. The model explains none of the variation

✓ **C. The model explains 100% of the variation in Y**

D. The slope coefficient equals one

Lecture 15: Multicollinearity

Q276. A high R^2 combined with all individually insignificant t-statistics is a classic symptom of:

A. Heteroskedasticity

✓ **B. Multicollinearity**

C. Autocorrelation

D. Missing data

Lecture 2: Summation Algebra & Linear Functions

Q277. Elasticity of demand measures:

A. The absolute change in quantity for a price change

✓ **B. The percentage change in quantity demanded for a one-percent change in price**

C. The total revenue change

D. The slope of the demand curve in original units

Lecture 11: Transformation & Functional Forms in Regression

Q278. In the regression $Y = \beta0 + \beta1X + \beta2X^2 + e$, the model is considered 'linear' because:

A. X only has power 1

✓ **B. The model is linear in the PARAMETERS $\beta0$, $\beta1$, $\beta2$**

C. The graph is a straight line

D. The squared term is not really a variable

Lecture 6: Simple Regression Model & OLS

Q279. In the regression context, the assumption 'E(e|X) = 0' means:

A. Errors are always positive

✓ **B. The expected value of the error, given any value of X, is zero**

C. X has no effect on Y

D. The intercept is zero

Lecture 5: Multivariate Optimization & Probability Review

Q280. Multivariate optimization requires setting ALL first-order partial derivatives equal to zero because:

A. Only one variable matters
✓ **B. A critical point must have zero slope in all directions simultaneously**
C. The Hessian is always positive
D. Cross partial derivatives are always zero

Lecture 7: Estimation, Testing & R^2 in Regression

Q281. The ANOVA table in regression output decomposes the Total Sum of Squares (TSS) into:

A. RSS + Error
✓ **B. Regression SS + Residual SS**
C. $R^2 + (1 - R^2)$
D. n + k

Lecture 6: Simple Regression Model & OLS

Q282. The key insight of OLS is to find estimates a and b that:

A. Maximise the sum of squared residuals
✓ **B. Minimise the sum of squared residuals**
C. Set all residuals equal to zero
D. Maximise the number of observations

Lecture 7: Estimation, Testing & R^2 in Regression

Q283. If P-value < α (level of significance) for a coefficient, we:

A. Fail to reject H0
✓ **B. Reject H0 and conclude the coefficient is statistically significant**
C. Conclude the model is a poor fit
D. Remove the variable from the model

Q284. The number of degrees of freedom for a simple regression (k = 2 parameters) with n = 20 observations is:

A. 20
B. 2
✓ **C. 18**
D. 10

Lecture 12: Standardised Variables & Regression Strategy

Q285. In a standardised regression, the intercept is always zero because:

A. The model is forced through the origin
✓ **B. The means of standardised X* and Y* are both zero**
C. The slope coefficient absorbs the intercept
D. Standardisation removes variation from the data

Lecture 9: Multiple Regression

Q286. Which formula correctly converts $\sum x1^2$ from normal to deviation form?

A. $\sum X1^2 - \sum X1$
✓ **B. $\sum X1^2 - (\sum X1)^2/n$**

C. $\sum X1^2 / n$
D. $\sum X1^2 + (\sum X1)^2/n$

Lecture 11: Transformation & Functional Forms in Regression

Q287. The primary reason we use natural logarithms in regression is that they can:

A. Increase the number of observations
✓ **B. Linearise certain non-linear relationships and reduce skewness**
C. Guarantee homoskedasticity
D. Eliminate all outliers

Lecture 7: Estimation, Testing & R^2 in Regression

Q288. A t-statistic of 6.51 compared to a critical value of t0.025,3 = 3.182 leads us to:

A. Accept H0: b = 0
✓ **B. Reject H0 and conclude b is statistically significant**
C. Conclude the model is a poor fit
D. Conclude heteroskedasticity exists

Lecture 9: Multiple Regression

Q289. If a regression has $R^2 = 0.9979$ and $N - k = 7$ and $k - 1 = 2$, and $F = 1667.53$, the conclusion of the F-test ($F\alpha = 4.74$) is:

A. The fit is not good
✓ **B. The fit is good**
C. R^2 is too low
D. The degrees of freedom are insufficient

Lecture 6: Simple Regression Model & OLS

Q290. The property that $\sum ei = 0$ after OLS estimation is a result of:

A. The assumption of normality
✓ **B. The first normal equation derived from minimising the sum of squared residuals**
C. The F-test condition
D. The Gauss-Markov theorem directly

Lecture 10: K-Variable Regression & Analysis ToolPak

Q291. Adding more independent variables always increases R^2, which is why we use:

A. Simple R^2 for model comparison
✓ **B. Adjusted R^2**
C. F-statistic only
D. The t-statistic only

Lecture 15: Multicollinearity

Q292. In a regression of GDP on Labour Force, GFCF, and Exports, a very high VIF for Exports (>10) suggests:

A. Exports does not belong in the model
✓ **B. Exports is collinear with other regressors**
C. GDP is explained solely by Exports
D. The F-test is invalid

Lecture 18: Heteroskedasticity: Formal Tests

Q293. The chi-square distribution is used in tests like the Park LM, Breusch-Pagan, and White tests because:

A. Residuals always follow a chi-square distribution
✓ **B. The LM statistic N × R^2 follows approximately a chi-square distribution under H0**
C. Chi-square is the same as the F-distribution
D. The original model errors are chi-square

Lecture 5: Multivariate Optimization & Probability Review

Q294. For profit maximisation $\pi = 80x - 2x^2 - xy - 3y^2 + 100y$, the critical values were x■ ≈ 16.52 and ■ ≈ 13.91. The second order condition for maximum requires:

A. $\pi xx > 0$ and $\pi yy > 0$
✓ **B. $\pi xx < 0$ and $\pi yy < 0$**
C. $\pi xx = 0$ and $\pi yy = 0$
D. $\pi xx > \pi yy$

Lecture 7: Estimation, Testing & R^2 in Regression

Q295. In a regression equation Y = 0.3 + 0.82X, the slope b = 0.82 means:

A. When X increases by 0.82, Y increases by 1
✓ **B. When X increases by 1 unit, Y increases by 0.82 units on average**
C. The correlation between X and Y is 0.82
D. Y equals 0.82 when X = 0

Lecture 15: Multicollinearity

Q296. The regression of a variable X1 on all other X variables to compute VIF is called:

A. The main regression
✓ **B. An Auxiliary Regression**
C. A Diagnostic Regression
D. The Secondary Model

Lecture 18: Heteroskedasticity: Formal Tests

Q297. The Breusch-Pagan test is sensitive to violations of normality. Which test is suggested as more robust?

A. Goldfeld-Quandt
✓ **B. White Test**
C. Park Test
D. Glesjer Test

Lecture 7: Estimation, Testing & R^2 in Regression

Q298. Which distribution is used to test the goodness of fit of the overall regression?

A. t-distribution
B. Normal distribution
✓ **C. F-distribution**
D. Chi-square

Lecture 1: Overview of Econometrics & Data Types

Q299. Econometrics is said to add 'empirical content' to economic theory. This means it:

A. Replaces theory with data
✓ **B. Allows theories to be tested against real-world data**
C. Makes economic theory purely mathematical
D. Eliminates the need for data collection

Lecture 5: Multivariate Optimization & Probability Review

Q300. When two coins are tossed, the probability of getting exactly one head is:

A. 1/4
B. 3/4
✓ **C. 1/2**
D. 1

Lecture 7: Estimation, Testing & R^2 in Regression

Q301. The degrees of freedom for the t-test of a slope coefficient in simple regression with n = 5 and 2 parameters is:

A. 5
B. 2
✓ **C. 3**
D. 1

Lecture 18: Heteroskedasticity: Formal Tests

Q302. In the Goldfeld-Quandt Test with N = 30, C = 6, and K = 2, the degrees of freedom for each sub-regression is:

✓ **A. 12**
B. 14
C. 10
D. 16

Lecture 1: Overview of Econometrics & Data Types

Q303. Which econometric software is introduced AFTER mid-term examination in ECO601?

A. SPSS
B. SAS
✓ **C. Stata**
D. EViews

Lecture 7: Estimation, Testing & R^2 in Regression

Q304. The intercept in a Keynesian consumption function represents:

A. Induced consumption
✓ **B. Autonomous consumption (consumption even when income is zero)**
C. Marginal propensity to consume
D. Total expenditure

Lecture 9: Multiple Regression

Q305. In multiple regression, the coefficient $\beta 2$ represents the effect of X2 on Y:

A. Without controlling for other variables

✓ **B. While holding all other variables constant (ceteris paribus)**
C. Only when X1 = 0
D. Only in simple regression

Lecture 6: Simple Regression Model & OLS

Q306. The assumption 'Random Sampling of n observations' in OLS ensures:

A. All observations are from the same date
✓ **B. The sample is representative of the population, allowing generalisation**
C. There is no error in the model
D. All X values are different

Lecture 10: K-Variable Regression & Analysis ToolPak

Q307. If the F-statistic from the Analysis ToolPak shows p = 0.0000000178, we conclude at 1% level that:

A. The fit is not good
✓ **B. The fit is good (reject H0 of no fit)**
C. R^2 is too low
D. The model has multicollinearity

Lecture 5: Multivariate Optimization & Probability Review

Q308. The mean of the Chi-Square distribution with k degrees of freedom equals:

A. k/2
B. 2k
✓ **C. k**
D. 1/k

Q309. The variance of the Chi-Square distribution equals:

A. k
B. k^2
✓ **C. 2k**
D. k/2

Lecture 6: Simple Regression Model & OLS

Q310. A 'Best' estimator (in the BLUE sense) has:

A. Zero bias
✓ **B. The smallest variance among all linear unbiased estimators of the parameter**
C. The largest sample size
D. The highest R^2

Lecture 18: Heteroskedasticity: Formal Tests

Q311. For the LM test statistics used in heteroskedasticity tests, LM = N × R^2 where R^2 and N are taken from:

A. The main regression
✓ **B. The auxiliary regression (regressing e^2 or $|e|$ or ln e^2 on the X variables)**
C. The scatter diagram
D. The Adjusted R^2 formula

Lecture 13: Qualitative Variables & Dummy Variables

Q312. Adding an interaction term G×E to a wage regression allows the:

A. Effect of gender G on wages to be tested
✓ **B. Effect of education E on wages to differ between males and females**
C. Intercept to remain constant across groups
D. Model to become non-linear in parameters

Lecture 5: Multivariate Optimization & Probability Review

Q313. The Probability Density Function (PDF) for a continuous random variable gives probability as:

A. A single point value
✓ **B. The area under the curve between two values (an integral)**
C. The height of the curve at a point
D. The sum of all possible values

Q314. The Cumulative Distribution Function (CDF) F(X) = P(X ≤ x) gives:

A. The probability that X equals exactly x
✓ **B. The probability that X is less than or equal to x**
C. The density at point x
D. The expected value of X

Q315. In the t-distribution, as the degrees of freedom increase toward infinity, the distribution approaches:

A. The Chi-square distribution
B. The F-distribution
✓ **C. The Standard Normal distribution**
D. The Uniform distribution

Lecture 15: Multicollinearity

Q316. Under perfect multicollinearity, the two normal equations effectively reduce to one because:

✓ **A. X1 and X2 carry the same information (linear combination)**
B. There is no variation in Y
C. The sample size is too small
D. The slope coefficients are both zero

Q317. Which of the following is an example of a 'Near' multicollinearity situation in economics?

A. GDP and GDP are the same variable
✓ **B. Imports and Exports both sharing a common upward trend in time series data**
C. Education and wages are perfectly correlated
D. Gender and race are perfectly correlated

Lecture 7: Estimation, Testing & R^2 in Regression

Q318. If $\sum e_i^2 = 11.9$ and degrees of freedom = 3, then the standard error of estimate σ_e is approximately:

A. 3.97
✓ **B. 1.99**
C. 4.57
D. 2.43

Q319. The LINEST Excel formula, when 'const = FALSE', forces:

A. R^2 to be zero
✓ **B. The intercept 'a' to be zero (regression through the origin)**
C. All coefficients to be positive
D. The slope to equal R^2

Lecture 18: Heteroskedasticity: Formal Tests

Q320. White's test auxiliary regression is sometimes simplified by regressing e^2 on:

A. X1 and X2 only
✓ **B. ■ and ■² (estimated Y and its square)**
C. All original X variables only without cross products
D. The original dependent variable Y

Lecture 15: Multicollinearity

Q321. A researcher finds high R^2 (0.99) and F = 1667 but $\beta 1$ is insignificant (t = –1.287). The likely problem is:

A. Heteroskedasticity
✓ **B. Multicollinearity ($\beta 1$ and $\beta 2$ are collinear)**
C. Autocorrelation
D. Wrong functional form

Lecture 6: Simple Regression Model & OLS

Q322. In OLS, the condition $\sum ei \times Xi = 0$ ensures that:

A. The errors are uncorrelated with each other
✓ **B. The estimated errors are uncorrelated with the explanatory variable**
C. The slope coefficient is zero
D. The intercept is zero

Lecture 5: Multivariate Optimization & Probability Review

Q323. A 'point of inflection' in multivariate optimisation is ruled out by the condition:

A. First order conditions equal zero
✓ **B. $Zxx \times Zyy > Zxy^2$**
C. $Zxx < 0$ and $Zyy < 0$
D. All partial derivatives are positive

Lecture 13: Qualitative Variables & Dummy Variables

Q324. If a dummy variable coefficient is not significant, it suggests:

A. The category captured by the dummy has a different intercept
✓ **B. The category captured by the dummy does NOT have a statistically different intercept from the reference category**
C. The reference category has no meaning
D. All categories are identical

Lecture 9: Multiple Regression

Q325. In the LINEST output for multiple regression with X1 and X2, the coefficients appear as:

A. $\beta 0$, $\beta 1$, $\beta 2$ from left to right

✓ **B. $\beta 2, \beta 1, \beta 0$ from left to right**

C. $\beta 1, \beta 2, \beta 0$ from left to right

D. Alphabetically

Lecture 15: Multicollinearity

Q326. If Tolerance = 0.002 for an explanatory variable, this indicates:

A. Mild multicollinearity (T is between 0.1 and 0.2)

✓ **B. Severe / Problematic multicollinearity ($T \leq 0.1$)**

C. No multicollinearity

D. Heteroskedasticity

Lecture 7: Estimation, Testing & R^2 in Regression

Q327. The standard error se(b) for the slope in simple regression involves which quantity in the denominator?

A. $\sqrt{(\sum X^2)}$

✓ **B. $\sqrt{[\sum X^2 - (\sum X)^2/n]}$**

C. n

D. $\sum e^2$

Lecture 18: Heteroskedasticity: Formal Tests

Q328. The White Test is more general than Breusch-Pagan because:

A. It is faster to compute

✓ **B. It includes squares and cross-products of regressors and does not require normality**

C. It uses the t-distribution

D. It assumes a linear relationship between variance and X

Lecture 7: Estimation, Testing & R^2 in Regression

Q329. For the simple regression line, the coefficient of determination R^2 equals the square of the correlation coefficient. This is true ONLY for:

A. Multiple regression

✓ **B. Simple regression (one independent variable)**

C. Log-log regression

D. Standardised regression

Lecture 18: Heteroskedasticity: Formal Tests

Q330. The degrees of freedom for the Chi-square distribution used in LM tests equal:

A. N – 1

✓ **B. K – 1 (number of regressors in auxiliary regression minus one)**

C. N – K

D. 1

Lecture 1: Overview of Econometrics & Data Types

Q331. In econometrics, 'model specification' refers to:

A. Collecting the data

✓ **B. Writing a mathematical equation to reflect the theoretical relationship among variables**

C. Estimating the model parameters

D. Publishing the results

Lecture 6: Simple Regression Model & OLS

Q332. OLS estimation produces residuals that sum to zero ($\sum ei = 0$) only when:

A. The model has no intercept

✓ **B. The model includes an intercept (constant) term**

C. The R^2 is greater than 0.9

D. There are more than 30 observations

Lecture 7: Estimation, Testing & R^2 in Regression

Q333. Expenditure on food is regressed on income. The regression equation Y = 0.3 + 0.82X is obtained. The trend value for X = 35 is:

A. 0.3

✓ **B. 29**

C. 0.82

D. 28.7

Lecture 13: Qualitative Variables & Dummy Variables

Q334. If the estimated wage gap between Black and White workers (reference) is –45.17 and is significant at 1%, this means:

A. Black workers earn 45.17 more than White workers

✓ **B. Black workers earn, on average, 45.17 rupees less per hour than White workers (same education, experience, gender)**

C. White workers earn 45.17 less than Black workers

D. Race has no effect on wages

Lecture 5: Multivariate Optimization & Probability Review

Q335. A variable whose value is determined by the outcome of a random experiment is called a:

A. Population parameter

✓ **B. Stochastic or Random Variable**

C. Sample statistic

D. Deterministic variable

Lecture 3: Quadratic, Exponential, Logarithmic Functions & Derivatives

Q336. The Price Elasticity of Demand formula using derivatives is:

A. Ep = dP/dQ × Q/P

✓ **B. Ep = dQ/dP × P/Q**

C. Ep = dQ/dP

D. Ep = dP/dQ

Lecture 1: Overview of Econometrics & Data Types

Q337. Econometrics adds 'empirical content' to economic theory. This can also be described as the process of:

A. Creating new theories without data

✓ **B. Testing and quantifying economic relationships using real-world data**

C. Replacing mathematical models with qualitative descriptions

D. Eliminating all uncertainty in economic prediction

Lecture 11: Transformation & Functional Forms in Regression

Q338. A non-linear relationship that can be LINEARISED by taking logs is called:

A. A polynomial model
✓ **B. A logarithmic model (intrinsically linear after transformation)**
C. A quadratic model
D. A dynamic model

Lecture 7: Estimation, Testing & R^2 in Regression

Q339. In the F-distribution, the numerator degrees of freedom (k−1) represents the number of:

A. Observations
✓ **B. Restrictions imposed or number of slope coefficients in regression**
C. Independent variables plus one
D. Parameters in the auxiliary regression

Lecture 6: Simple Regression Model & OLS

Q340. Which of the following best describes the term 'Regressand'?

A. The independent variable in regression
✓ **B. The dependent variable (the variable we are explaining)**
C. The error term
D. The slope coefficient

Lecture 3: Quadratic, Exponential, Logarithmic Functions & Derivatives

Q341. For a quadratic demand equation $Qd = aP^2 + bP + c$, if we want to find price that minimises (or maximises) quantity demanded, we use:

A. Integration
✓ **B. Setting $dQd/dP = 0$ (First Order Condition)**
C. Setting $d^2Qd/dP^2 = 0$
D. The Sum-Difference Rule

Lecture 15: Multicollinearity

Q342. The R^2 in the auxiliary regression of X1 on X2 and X3 was 0.998. The corresponding VIF equals:

✓ **A. 500**
B. 0.002
C. 2
D. 998

Lecture 17: Heteroskedasticity: Causes & Detection

Q343. Under Heteroskedasticity, estimates of variance are biased, which means:

A. Slope coefficients are biased
✓ **B. t and F tests are invalid because they are based on incorrect variance estimates**
C. R^2 will be negative
D. The number of observations is reduced

Lecture 18: Heteroskedasticity: Formal Tests

Q344. In the White test, running $e^2 = \gamma 0 + \gamma 1 \blacksquare + \gamma 2 \blacksquare^2 + v$ is an alternative to the main auxiliary regression because:

A. It is always more powerful
✓ **B. It avoids having too many regressors when K is large (uses fitted values as a proxy)**
C. It does not require any additional computation
D. It tests for multicollinearity instead

Lecture 1: Overview of Econometrics & Data Types

Q345. Which of the following best describes Econometrics according to its modern view?

A. A subject limited to economic theory testing
✓ **B. Application of math, statistics, and computer science to analyse quantitative OR qualitative phenomena from any discipline**
C. A branch of pure mathematics
D. A method used only by central banks

Q346. The step 'Model Estimation' in the econometric methodology involves:

A. Writing the hypothesis
✓ **B. Using an appropriate econometric technique to find the values of the parameters**
C. Collecting the data
D. Publishing forecasts

Q347. Which is an example of an economic question that Econometrics can answer?

A. What is 2 + 2?
✓ **B. Are oil prices and the value of the US dollar correlated?**
C. What colour is the sky?
D. How many planets orbit the sun?

Q348. The phrase 'From Idea to Forecasting' in the methodology summarises how many phases?

A. 2
B. 3
✓ **C. 5**
D. 7

Q349. Penn World Tables is a data source for:

A. Medical research
✓ **B. Macroeconomic data across countries**
C. Literary analysis
D. Weather forecasting

Lecture 2: Summation Algebra & Linear Functions

Q350. Which of the following correctly states Property 3 of summation?

A. $\sum(Xi + Yi) = \sum Xi \times \sum Yi$
✓ **B. $\sum(Xi + Yi) = \sum Xi + \sum Yi$**
C. $\sum(Xi + Yi) = \sum Xi - \sum Yi$
D. $\sum(Xi + Yi) = (\sum Xi + \sum Yi)/n$

Q351. For the linear demand function $Qd = 50 - 2P$, what is the slope of the demand curve?

A. 50
✓ **B. -2**
C. 2

D. 25

Q352. If n = 5 and a = 3, then ∑a (summed from i=1 to 5) equals:

A. 3
B. 5
✓ C. 15
D. 8

Q353. In a linear function Y = 5 + 2X, the slope b = 2 means:

A. Y equals 2 when X = 0
✓ B. For every one unit increase in X, Y increases by 2 units
C. Y increases by 5 when X increases by 1
D. The function passes through the origin

Q354. Expanding the expression ∑(k=2 to 6) k/(k+1) gives:

✓ A. 2/3 + 3/4 + 4/5 + 5/6 + 6/7
B. 2/3 + 3/4 + 4/5
C. 1/2 + 2/3 + 3/4
D. 6/7 + 5/6 + 4/5 + 3/4 + 2/3 (reversed)

Q355. The double summation is used when:

A. We need to sum over a single subscript
✓ B. We sum over two subscripts simultaneously (nested summation)
C. We square the entire sum
D. We divide the total by n

Q356. The extension of summation property 2 and 3 gives: ∑(aXi + bYi) =

A. ab × ∑(XiYi)
✓ B. a∑Xi + b∑Yi
C. (a+b)∑Xi
D. ∑Xi × ∑Yi

Lecture 3: Quadratic, Exponential, Logarithmic Functions & Derivatives

Q357. A quadratic function Y = aX² + bX + c has how many parameters?

A. 1
B. 2
✓ C. 3
D. 4

Q358. The graph of a quadratic function is called a:

A. Hyperbola
✓ B. Parabola
C. Ellipse
D. Straight line

Q359. For the quadratic function Y = 2X² + 5, when does the graph reach its minimum?

A. At X = 5
B. At X = 2
✓ C. At X = 0 (vertex at c=5)
D. At X = –2

Q360. The Kuznets curve is an example of which type of curve?

A. U-shaped (minimum)
✓ **B. Inverted-U shaped (maximum)**
C. Straight line
D. Exponential

Q361. Euler's constant 'e' is the base of the:

A. Common logarithm (base 10)
✓ **B. Natural logarithm (base e)**
C. Binary logarithm (base 2)
D. Quadratic logarithm

Q362. The derivative of the exponential function Y = e■ is:

A. e■■¹
✓ **B. e■**
C. xe^(x-1)
D. 1/e■

Q363. For the chain rule, if Y = (5x² + 2x + 10)³, then dY/dX equals:

A. 3(5x² + 2x + 10)²
✓ **B. 3(5x² + 2x + 10)²(10x + 2)**
C. (5x² + 2x + 10)³ × (10x + 2)
D. 3 × (10x + 2)

Q364. The Sum-Difference Rule of differentiation states that d/dx[f(x) ± g(x)] =

A. f(x) × g'(x) ± g(x) × f'(x)
✓ **B. f'(x) ± g'(x)**
C. f'(x) × g'(x)
D. f(x)/g'(x) ± g(x)/f'(x)

Lecture 4: Partial Derivatives & Simple Optimization

Q365. For Z = 2x²y² + 5x³y■, the cross partial Zxy equals:

A. 4xy² + 15x²y■
B. 4x²y + 20x³y³
✓ **C. 8xy + 60x²y³**
D. 4y² + 30xy■

Q366. The notation Zxx means:

A. ∂Z/∂x
✓ **B. The second partial derivative of Z differentiated twice with respect to x**
C. The cross partial ∂²Z/(∂x∂y)
D. The partial derivative with respect to y

Q367. When we differentiate Z = ∑(y – a – bx)² with respect to 'a' and set equal to zero, we use the:

A. Power rule only
✓ **B. Chain rule (applied to the inner expression) and summation algebra**
C. Quotient rule only
D. Product rule only

Q368. A function Qd = f(P, Y, Ps, Pc, A) has how many first-order partial derivatives?

A. 1

B. 3

✓ **C. 5**

D. 7

Q369. For the profit function $\pi = 100Q - 120 - 2Q^2$, the second derivative $\pi'' =$

A. 100 – 4Q

✓ **B. –4**

C. 100

D. 4

Lecture 5: Multivariate Optimization & Probability Review

Q370. When the Hessian determinant Zxx × Zyy – Zxy² < 0, the critical point is a:

A. Local minimum

B. Local maximum

✓ **C. Saddle point**

D. Global maximum

Q371. Collectively Exhaustive Events exhaust all possible outcomes. For a single coin toss:

A. Only heads exhausts all outcomes

✓ **B. Heads and tails together are collectively exhaustive**

C. Neither heads nor tails is collectively exhaustive

D. The event must contain more than two outcomes

Q372. Equally Likely Events means:

A. One event prevents the other

✓ **B. Each event is as likely as the other to occur**

C. All events are mutually exclusive

D. Events cannot be both equally likely and mutually exclusive

Q373. The probability mass function f(X = xi) satisfies which condition?

A. $\sum f(xi) = 0$

✓ **B. $\sum f(xi) = 1$**

C. f(xi) can be negative

D. f(xi) > 1 for some values

Q374. For the F-distribution, the denominator degrees of freedom k2 corresponds to:

A. The sample size

✓ **B. The degrees of freedom of the denominator Chi-square**

C. The number of independent variables

D. The level of significance

Lecture 6: Simple Regression Model & OLS

Q375. In regression, the notation ■ (Y-hat) represents:

A. The actual observed value of Y

✓ **B. The estimated (fitted) value of Y from the regression**

C. The error term

D. The mean of Y

Q376. The regression residual ei is computed as:

A. Yi – X■

✓ **B. Yi – ■i**

C. ■i – Yi

D. Yi – β0

Q377. The minimisation of ∑ei² with respect to 'a' gives which normal equation?

A. ∑XY = a∑X + b∑X²

✓ **B. ∑Y = na + b∑X**

C. ∑e = 0

D. ∑X = na + b∑Y

Q378. Under Gauss-Markov assumptions, OLS produces BLUE estimators. 'U' stands for:

A. Uniform

✓ **B. Unbiased**

C. Unique

D. Uncorrelated

Q379. The term 'Conditional Zero Mean of Errors' means:

A. E(e) = 0 unconditionally

✓ **B. E(e|X) = 0 for all values of X**

C. Errors are always positive

D. Errors have zero correlation with Y

Lecture 7: Estimation, Testing & R² in Regression

Q380. The formula for computing R² using sums is:

✓ **A. R² = 1 – N∑e² / (N∑Y² – (∑Y)²)**

B. R² = N∑e² / ∑Y²

C. R² = ∑e² / N

D. R² = (∑e)² / ∑Y²

Q381. A t-value of 2.5 with tα/2, n-k = 2.306 at α = 0.05 leads to the conclusion:

A. b is not significant at the 5% level

✓ **B. b is significant at the 5% level (reject H0)**

C. The model has a poor fit

D. R² is greater than 0.9

Q382. The formula for F-statistic from R² is:

A. F = R² / (1 – R²)

✓ **B. F = [R²/(k–1)] / [(1–R²)/(N–k)]**

C. F = R² × N

D. F = (1–R²) × (k–1)

Q383. If R² = 0 in simple regression, it means:

A. Perfect fit

✓ **B. X explains none of the variation in Y (X is useless as a predictor)**

C. All residuals are zero

D. b = 1

Q384. The term 'Standard Error of 'b'' measures:

A. How far b is from the true β0

✓ **B. The sampling variability (precision) of the slope estimate b**

C. The variance of Y

D. The variance of e

Lecture 8: Excel Applications in Regression

Q385. In the transformation $Y = AK^{\beta} \rightarrow \ln Y = \ln A + \beta \ln K$, after regression, to get the original parameter A you:

A. Multiply the estimated lnA by e
✓ B. Take the antilog (exponentiate) of the estimated intercept: A = e^(estimated intercept)
C. Square the estimated intercept
D. Divide by β

Q386. Which formula in Excel computes R-Squared for a simple regression?

✓ A. =RSQ(Y range, X range)
B. =CORREL(Y range, X range)
C. =VAR(Y range)
D. =COVAR(Y range, X range)

Q387. The phrase 'Change of unit does not change correlation' means that if we change the scale of X or Y:

A. The regression slope stays the same
✓ B. The linear correlation coefficient r remains unchanged
C. R^2 increases
D. The intercept becomes zero

Q388. An outlier that is an 'extreme X value' is most likely to:

A. Have a zero residual
✓ B. Affect the slope of the regression line significantly
C. Make R^2 equal to zero
D. Remove multicollinearity

Q389. In OLS regression, 'Sum of Squared Residuals' is also called:

A. SSR (Sum of Squares Regression)
✓ B. SSE (Sum of Squares Error or Residual Sum of Squares)
C. SST (Sum of Squares Total)
D. R-squared

Lecture 9: Multiple Regression

Q390. When we say we are 'isolating the unique effect' of one variable in multiple regression, we mean:

A. We run a simple regression with that variable only
✓ B. We hold all other variables constant (ceteris paribus), capturing the partial effect
C. We remove all other variables
D. We standardise only that variable

Q391. In multiple regression, the formula for $\beta 2$ (in deviation form) has the same structure as $\beta 1$ but with x1 and x2:

A. Added together
✓ B. Swapped (symmetric structure)
C. Divided by each other
D. Multiplied together

Q392. The normal equations for multiple regression with 2 independent variables are derived by minimising:

A. $\sum Yi$

✓ **B. $\sum(Y - \beta 0 - \beta 1X1 - \beta 2X2)^2$**

C. $\sum(X1i + X2i)$

D. $\sum ei \times Yi$

Q393. The formula $\sum x2y = \sum X2Y - (\sum X2 \sum Y)/n$ converts summation from normal to deviation form for:

✓ **A. $\sum x2y$**

B. $\sum x2^2$

C. $\sum x1x2$

D. $\sum y^2$

Q394. In the multiple regression example with 10 observations, $\beta 2 = -5.1713$ means:

A. For every unit increase in X2, Y increases by 5.1713

✓ **B. For every unit increase in X2 (holding X1 constant), Y decreases by 5.1713 on average**

C. X2 is not significant

D. The model is misspecified

Lecture 10: K-Variable Regression & Analysis ToolPak

Q395. To install the Data Analysis ToolPak in Excel 2007, you go to:

A. Insert tab → Add-Ins

✓ **B. Office Button → Excel Options → Add-Ins → Manage Excel Add-Ins → GO**

C. View tab → Analysis

D. Home tab → Analysis

Q396. When the Data Analysis ToolPak is successfully installed, you see 'Data Analysis' under the:

A. Home ribbon

✓ **B. DATA ribbon (right-most item)**

C. Insert ribbon

D. View ribbon

Q397. In the ANOVA section of Analysis ToolPak regression output, 'SS/df = MS' where MS stands for:

A. Mean Sum

✓ **B. Mean Square**

C. Minimum Square

D. Multiple Statistics

Q398. The 'Significance F' in the Analysis ToolPak output is the p-value for:

A. Individual t-tests

✓ **B. The overall F-test (goodness of fit)**

C. The intercept only

D. R-squared

Q399. In the Analysis ToolPak regression output, 'Lower 95%' and 'Upper 95%' represent:

A. The range of Y values

✓ **B. The 95% confidence interval for each coefficient**

C. The range of X values

D. The t-statistic bounds

Lecture 11: Transformation & Functional Forms in Regression

Q400. In the log-log model (double-log) lnY = a + b lnX, the marginal effect dY/dX equals:

A. b
✓ **B. b(Y/X)**
C. b/X
D. b × Y

Q401. Which functional form implies that each additional unit of X has a SMALLER and SMALLER effect on Y?

A. Y = a + bX (linear)
✓ **B. Y = a + b lnX (level-log or logarithmic form)**
C. Y = a + bX² (quadratic)
D. lnY = a + bX (log-level)

Q402. A polynomial regression Y = β0 + β1X + β2X² + β3X³ + e can STILL be estimated by OLS because:

A. The variables have only power 1
✓ **B. The model is linear in PARAMETERS β0, β1, β2, β3**
C. Taking logs eliminates the polynomial terms
D. The cubic term is ignored in OLS

Q403. In a model Y = a + bI + cH + d(I×H) + e, the marginal effect of H on Y is:

A. c
✓ **B. c + dI**
C. d
D. c + d

Q404. An autoregressive model Yt = a + b Yt–1 + e uses:

✓ **A. The lagged value of the DEPENDENT variable as a regressor**
B. The lagged value of the INDEPENDENT variable
C. The squared value of Y
D. The first difference of Y

Lecture 12: Standardised Variables & Regression Strategy

Q405. The advantage of Beta (standardised) coefficients in comparing predictors is that they:

A. Are always larger than unstandardised coefficients
✓ **B. Allow comparison of effects across variables with different units**
C. Always equal zero
D. Are the same as elasticities

Q406. Before performing regression, computing SUMMARY STATISTICS helps because:

A. It estimates the regression coefficients
✓ **B. It gives insight into the distribution, spread, and central tendency of variables**
C. It eliminates multicollinearity
D. It automatically corrects missing data

Q407. In Excel's Data Analysis ToolPak, to get Summary Statistics you select:

A. Regression
✓ **B. Descriptive Statistics**

C. Correlation
D. Histogram

Q408. A HISTOGRAM created before regression is used to:

A. Check R^2
✓ B. Visualise the distribution of a variable (check for skewness, outliers)
C. Estimate the slope coefficient
D. Test for multicollinearity

Q409. When variables in the same model measure similar things, one option is to:

A. Include all of them to increase R^2
✓ B. Combine them into a single composite variable or keep only one
C. Use dummy variables for each
D. Take the log of each separately

Lecture 13: Qualitative Variables & Dummy Variables

Q410. If we assign G=1 for males and G=0 for females, the interpretation of $\beta 3$ changes to:

A. $\beta 3$ now captures the female-male wage gap
✓ B. $\beta 3$ now captures the male-female wage gap (premium for males)
C. $\beta 3$ remains unchanged in meaning
D. $\beta 3$ becomes zero

Q411. When all three race dummies (White, Black, Asian) are included in a model, the problem is:

A. Heteroskedasticity
✓ B. Perfect multicollinearity (Dummy Variable Trap)
C. Omitted variable bias
D. Autocorrelation

Q412. For a polytomous qualitative variable with 5 categories, how many dummies are needed in regression?

A. 5
✓ B. 4
C. 3
D. 6

Q413. If $\beta 4$ (coefficient on interaction G×D1) is significant and negative, it means:

A. Black males earn more than White males
✓ B. Black women earn significantly less than Black men (additional gap beyond Black-White and male-female gaps)
C. The model is incorrect
D. The reference is Black males

Q414. Dummy variables can also capture STRUCTURAL BREAKS in time series data, for example:

✓ A. A dummy = 1 after 2008 financial crisis captures a shift in the intercept
B. A dummy = 1 for all observations increases R^2
C. A dummy = 1 for large firms makes the model non-linear
D. A dummy = 1 eliminates heteroskedasticity

Lecture 14: Logarithms in Regression & Missing Data

Q415. When using log(Wage) instead of Wage in regression, R^2 improved because:

A. Log transformation reduces the number of observations
✓ **B. Log transformation reduces right skewness and better satisfies regression assumptions**
C. Log values are always larger than original values
D. Log transformation removes all heteroskedasticity automatically

Q416. Attrition in panel studies occurs when:

A. New observations are added each period
✓ **B. Respondents drop out over time (death, migration, dropout) creating missing data**
C. Variables are added to the model
D. The model is re-estimated each year

Q417. Which missing data mechanism is MOST BENIGN (easiest to handle)?

A. MNAR
B. MAR
✓ **C. MCAR**
D. None of them

Q418. Conditional questions in surveys (e.g., 'If employed, what is your income?') create which type of missing data?

A. Missing Completely At Random
✓ **B. Structured missing data — MAR or MNAR depending on the context**
C. Listwise deletion
D. Multiple imputation

Q419. Regression imputation to replace missing values tends to:

A. Increase variability in the data
✓ **B. Overestimate model fit and weaken variance (underestimating true variance)**
C. Eliminate all bias
D. Produce unbiased standard errors

Lecture 15: Multicollinearity

Q420. Multicollinearity makes it DIFFICULT to:

A. Compute R^2
✓ **B. Assess the individual contribution of each regressor to explained variation**
C. Run the regression at all
D. Calculate the intercept

Q421. Wide confidence intervals in regression are a consequence of:

A. Small standard errors (due to multicollinearity)
✓ **B. Large standard errors (due to multicollinearity increasing variance of estimators)**
C. Large sample size
D. High R^2

Q422. The correlation matrix in the multicollinearity example showed correlation between income and price of sub-1 as 0.7936. This suggests:

A. No multicollinearity between these variables
✓ **B. Potential multicollinearity between income and price of sub-1**
C. Perfect multicollinearity
D. Heteroskedasticity

Q423. Multicollinearity being a 'sample phenomenon' means:

A. It is always present in every dataset

✓ **B. It may exist in the particular sample even if not in the population — it depends on the data collected**

C. It is caused by sample size being too large

D. It disappears when we use the entire population

Q424. An example of a 'Natural Constraint' causing multicollinearity is:

A. Small sample size

✓ **B. In a country, high imports typically accompany high exports — both tend to move together**

C. Using dummy variables

D. Adding polynomial terms to the model

Lecture 16: Multicollinearity: Remedial Measures

Q425. After combining exports and imports into 'openness' (volume of trade) in the Pakistan GDP example, VIF:

A. Remained above 10

✓ **B. Decreased to around 5.36, indicating reduced multicollinearity**

C. Increased to 500

D. Became undefined

Q426. Transforming import and export data to 'real' values by dividing by CPI is an example of:

A. Dummy variable approach

✓ **B. Variable transformation to reduce multicollinearity**

C. Ridge regression

D. Listwise deletion

Q427. When Multicollinearity exists and the purpose is FORECASTING only (not causal inference), the recommendation is often:

A. Drop all correlated variables

✓ **B. Do nothing — forecasting can still be accurate even with multicollinearity**

C. Use the White test

D. Increase the level of significance

Q428. In the first difference remedial form ($Y_t - Y_{t-1}$), a limitation is that:

A. All coefficients become zero

✓ **B. One observation is lost, and the approach may not make sense for cross-sectional data**

C. R^2 always equals 1

D. The intercept becomes negative

Q429. Why might dropping a variable to remedy Multicollinearity cause problems?

A. Because R^2 increases artificially

✓ **B. Because dropping a theoretically relevant variable introduces omitted variable bias**

C. Because standard errors decrease

D. Because the F-test becomes invalid

Lecture 17: Heteroskedasticity: Causes & Detection

Q430. In Heteroskedasticity, OLS gives MORE weight to extreme values (large residuals) because:

A. OLS minimises sum of absolute residuals

✓ **B. OLS minimises sum of SQUARED residuals, so large errors carry disproportionate weight**

C. OLS maximises R^2
D. OLS ignores small residuals

Q431. 'Skedasis' in Heteroskedasticity means:

A. Same
B. Regression
✓ **C. Dispersion**
D. Error

Q432. Cross-country data often exhibits Heteroskedasticity because:

A. All countries report data perfectly
✓ **B. Different countries may have very different qualities and scales of data**
C. Time series data is always homoskedastic
D. Panel data is always heteroskedastic

Q433. If the variance of the error term Var(ei|X) increases with X, the pattern seen in residuals vs X plot would be:

A. Random scatter with constant spread
✓ **B. A funnel shape — spread widens as X increases**
C. A decreasing trend
D. No visible pattern

Q434. The consequence that 'OLS gives incorrect inference' under Heteroskedasticity means:

A. OLS cannot compute slope coefficients
✓ **B. t-statistics and F-statistics are invalid because they rely on biased variance estimates**
C. R^2 is biased downward
D. OLS produces biased intercepts

Lecture 18: Heteroskedasticity: Formal Tests

Q435. In the Goldfeld-Quandt Test, we reject H0 (Homoskedasticity) if:

A. F < Fα, n-c-k, n-c-k
✓ **B. F > Fα, (N-C)/2-K, (N-C)/2-K**
C. LM > χ^2
D. t > tα/2, n-k

Q436. The Breusch-Pagan test statistic LM = N × R^2 follows which distribution under H0?

A. t-distribution with n-k df
B. F-distribution
✓ **C. Chi-square distribution with K–1 degrees of freedom**
D. Normal distribution

Q437. In the Harvey-Godfrey test, the auxiliary regression is:

A. $e^2 = \alpha 0 + \alpha 1 \ln X1 + ...$
✓ **B. $\ln e^2 = \alpha 0 + \alpha 1 X1 + \alpha 2 X2 + ...$**
C. $|e| = \alpha 0 + \alpha 1 X1$
D. $Y = \alpha 0 + \alpha 1 e^2 + ...$

Q438. The Glesjer test differs from Breusch-Pagan in that it regresses:

A. Squared residuals on X variables
✓ **B. Absolute residuals on various transformations of X**
C. Log of squared residuals on log of X

D. Fitted Y on squared residuals

Q439. If the White test F-statistic from the auxiliary regression is significant, we conclude:

A. The model has multicollinearity
✓ **B. There is evidence of Heteroskedasticity**
C. The model has a perfect fit
D. The errors are normally distributed

Lecture 5: Multivariate Optimization & Probability Review

Q440. Which of the following is an implication of the Central Limit Theorem relevant to econometrics?

A. Small samples always have non-normal distributions
✓ **B. With large enough samples, the sampling distribution of sample means approaches normal**
C. Normal distribution applies only to discrete data
D. Econometric tests do not require any distributional assumption

Lecture 11: Transformation & Functional Forms in Regression

Q441. Choosing a log-log model for demand estimation (lnQd = a + b lnP) is appropriate because:

✓ **A. The coefficient b directly gives the price elasticity of demand**
B. The log transformation always gives a positive slope
C. Price and quantity are always normally distributed
D. It eliminates the need for an error term

Lecture 15: Multicollinearity

Q442. If X■ = 5X■ (perfect linear relationship) in a regression, the OLS estimation fails because:

A. The first normal equation has no solution
✓ **B. The normal equations become linearly dependent and cannot be uniquely solved**
C. The error term becomes zero
D. R^2 is automatically 1

Lecture 7: Estimation, Testing & R^2 in Regression

Q443. In testing H0: b = 0 vs H1: b ≠ 0 at α = 0.05, if the p-value = 0.03, we:

A. Fail to reject H0
✓ **B. Reject H0 at 5% and conclude b is significant**
C. Reject H0 at 1% as well
D. Cannot draw any conclusion

Lecture 8: Excel Applications in Regression

Q444. If β1 is the output elasticity with respect to labour in a Cobb-Douglas function, and β1 + β2 = 1, this implies:

A. Decreasing returns to scale
✓ **B. Constant returns to scale**
C. Increasing returns to scale
D. Zero returns to scale

Lecture 4: Partial Derivatives & Simple Optimization

Q445. The formula for partial derivative ensures we treat other variables as constants. This mirrors the economic concept of:

✓ **A. Ceteris paribus (other things equal)**
B. Mutatis mutandis
C. Pareto efficiency
D. Comparative statics (general equilibrium)

Lecture 6: Simple Regression Model & OLS

Q446. When we say OLS is 'Best' among linear unbiased estimators, we mean:

A. It gives the largest possible coefficient
✓ **B. It has the minimum variance among all linear unbiased estimators**
C. It is the fastest computational method
D. It gives exactly the true parameter

Lecture 7: Estimation, Testing & R^2 in Regression

Q447. A 95% confidence interval for the slope coefficient b means:

A. There is a 95% probability that the true $\beta 1$ is in this interval for this specific sample
✓ **B. In repeated sampling, 95% of such intervals constructed would contain the true $\beta 1$**
C. The slope is significant at the 5% level
D. R^2 is at least 95%

Lecture 15: Multicollinearity

Q448. A coefficient with an 'unexpected sign' in regression (e.g., positive slope for price in demand) suggests:

A. The model is correct and demand theory is wrong
✓ **B. Possible misspecification, omitted variables, or multicollinearity**
C. The variable should be dropped immediately
D. R^2 is too high

Lecture 7: Estimation, Testing & R^2 in Regression

Q449. In the simple regression example (food expenditure on income), a = 0.3 means:

A. When income is one unit, food expenditure is 0.3
✓ **B. At zero income, estimated food expenditure is 0.3 (the base/autonomous level)**
C. The slope is 0.3
D. R^2 is 0.3

Lecture 5: Multivariate Optimization & Probability Review

Q450. The Cumulative Distribution Function (CDF) F(X) is NON-DECREASING because:

A. Probability can be negative
✓ **B. Adding more area under the PDF always adds non-negative probability**
C. The PDF is always negative
D. X can only take increasing values

Lecture 13: Qualitative Variables & Dummy Variables

Q451. In the wage regression example with 35 observations, the $R^2 = 0.85$ means:

A. 15% of wages are unexplained

✓ **B. 85% of the variation in wages is explained by the model**

C. The F-statistic equals 0.85

D. There are 85 observations

Lecture 9: Multiple Regression

Q452. The notation β2 in the LINEST output appears in which position?

A. Rightmost cell of first row

✓ **B. Leftmost cell of first row**

C. Middle of first row

D. Second row, first cell

Lecture 11: Transformation & Functional Forms in Regression

Q453. For the model lnY = β0 + β1X (log-level), the elasticity at point (X0, Y0) is:

A. β1

✓ **B. β1 × X0**

C. β1 × Y0

D. β1 × X0 × Y0

Lecture 18: Heteroskedasticity: Formal Tests

Q454. The White Test is preferred over Breusch-Pagan when:

A. The sample is very small

✓ **B. The normality of errors cannot be assumed and the form of HSK is unknown**

C. There is only one independent variable

D. The model has no intercept

Lecture 5: Multivariate Optimization & Probability Review

Q455. The Chi-square distribution property 'mean equals k' where k = degrees of freedom means:

✓ **A. The average squared standard normal variable over k repetitions equals k**

B. The standard deviation of chi-square equals k

C. The mode of chi-square equals k – 2 for k > 2

D. The range is from –k to +k

Lecture 6: Simple Regression Model & OLS

Q456. For the OLS condition ∑eiXi = 0, this results from:

✓ **A. Setting the derivative of ∑e² with respect to b equal to zero**

B. Setting the derivative of ∑e² with respect to a equal to zero

C. The assumption of zero mean errors

D. The normality assumption

Lecture 12: Standardised Variables & Regression Strategy

Q457. In a regression table showing * for a coefficient, this typically means:**

A. The coefficient is NOT significant

✓ **B. The coefficient is significant at the 1% level**

C. The coefficient is significant at the 5% but not 1% level

D. The standard error is very large

Lecture 1: Overview of Econometrics & Data Types

Q458. Econometric methodology is described as going from 'Idea' to 'Forecasting'. The step between Model Estimation and Prediction is:

A. Data Collection
✓ **B. Performing Diagnostic Tests and Testing the Hypothesis**
C. Model Specification
D. Publication

Lecture 6: Simple Regression Model & OLS

Q459. When ∑(Xi − X■)² = 0, it implies that all X values are identical and OLS cannot be estimated.

A. False — OLS can still be estimated
✓ **B. True — there is no variation in X so the slope b is undefined**
C. Only true for time series
D. Only true for large samples

Q460. The 'Trend Value' ■i in regression is:

A. The actual observed Y
✓ **B. The best estimate of the expected value of Y given the observed X**
C. The maximum possible value of Y
D. The error in the regression

Lecture 16: Multicollinearity: Remedial Measures

Q461. Increasing sample size in cross-country regression can reduce Multicollinearity by:

A. Changing the model specification
✓ **B. Introducing more data variation that breaks the near-linear relationship between regressors**
C. Automatically dropping collinear variables
D. Reducing R^2

Lecture 17: Heteroskedasticity: Causes & Detection

Q462. A fan-shaped scatter of residuals vs X is indicative of:

A. Homoskedasticity
✓ **B. Heteroskedasticity (variance increasing with X)**
C. Perfect model fit
D. Multicollinearity

Lecture 7: Estimation, Testing & R^2 in Regression

Q463. For a regression with k = 3 parameters and N = 25, the degrees of freedom for t-test of coefficients are:

A. 25
B. 3
✓ **C. 22**
D. 8

Lecture 9: Multiple Regression

Q464. The formula for converting ∑x1x2 from normal to deviation form is:

A. ∑X1X2 + (∑X1∑X2)/n

✓ **B. $\sum X1X2 - (\sum X1 \sum X2)/n$**

C. $\sum X1X2 / n$

D. $n \times \sum X1X2$

Lecture 7: Estimation, Testing & R^2 in Regression

Q465. If r = –0.967 between X and Y in simple regression, then $R^2 \approx$

A. –0.967

✓ **B. 0.935**

C. 0.967

D. –0.935

Lecture 6: Simple Regression Model & OLS

Q466. In econometrics, the error term 'e' primarily represents:

A. Only the measurement error of X

✓ **B. Combined effect of all omitted variables plus random variability**

C. The coefficient estimation error

D. The difference between $\beta 0$ and $\beta 1$

Lecture 10: K-Variable Regression & Analysis ToolPak

Q467. The Input Y Range in the Analysis ToolPak Regression dialog should contain:

A. All X variables

✓ **B. The dependent variable values only**

C. Both X and Y values

D. The residuals only

Lecture 18: Heteroskedasticity: Formal Tests

Q468. The White Test auxiliary regression for 2 independent variables (X1, X2) includes:

A. X1 and X2 only

✓ **B. X1, X2, $X1^2$, $X2^2$, and X1×X2**

C. Only $X1^2$ and $X2^2$

D. e^2 and e as regressors

Lecture 9: Multiple Regression

Q469. For $\beta 1 = -0.0372$ with $se(\beta 1) = 0.02892$, the t-statistic is approximately:

A. 0.78

✓ **B. –1.29**

C. 1.29

D. –0.78

Lecture 7: Estimation, Testing & R^2 in Regression

Q470. The 'Total Variation' in Y (TSS) is decomposed into:

✓ **A. Explained Variation (ESS) + Residual Variation (RSS)**

B. $n \times R^2 + (1 - R^2)$

C. $SST = SSR + SSE + R^2$

D. $\sum Y^2 - \sum e^2$

Lecture 10: K-Variable Regression & Analysis ToolPak

Q471. If Adjusted R^2 decreases after adding a new variable, it suggests:

A. The new variable improved the model significantly
✓ **B. The variable did NOT improve fit enough to compensate for the lost degree of freedom**
C. The model is now perfect
D. Multicollinearity was eliminated

Lecture 11: Transformation & Functional Forms in Regression

Q472. Which model is NOT linear in parameters?

A. $Y = \beta0 + \beta1X + \beta2X^2$
B. $Y = \beta0 + \beta1\ln X$
✓ **C. $Y = \beta0 \times \beta1^X$**
D. $\ln Y = \beta0 + \beta1X$

Q473. An interaction term differs from correlation because:

A. Interaction means variables are correlated
✓ **B. Interaction means the EFFECT of one variable on Y depends on another variable — can exist without correlation**
C. Correlation captures joint effects while interaction does not
D. They are identical concepts

Lecture 17: Heteroskedasticity: Causes & Detection

Q474. Under Heteroskedasticity, standard OLS standard errors are:

A. Exactly correct
✓ **B. Different from true values, making t and F tests incorrect**
C. Always smaller than actual
D. Equal to zero

Lecture 6: Simple Regression Model & OLS

Q475. Ordinary Least Squares was popularised by:

A. Wooldridge
✓ **B. Gauss and Legendre**
C. Gujarati
D. Keynes

Lecture 9: Multiple Regression

Q476. When testing H0: $\beta2 = 0$ and $|t| < t\alpha/2,n-k$, we conclude:

A. $\beta2$ is significant
✓ **B. $\beta2$ is NOT significantly different from zero (fail to reject H0)**
C. The model is a poor fit
D. Multicollinearity exists

Lecture 13: Qualitative Variables & Dummy Variables

Q477. The 'reference category' in a dummy variable regression is:

A. The most important category
✓ **B. The baseline category assigned all-zero dummy values — the comparison group**
C. The category with highest mean

D. The first alphabetical category

Lecture 17: Heteroskedasticity: Causes & Detection

Q478. The Gauss-Markov theorem's BLUE guarantee breaks down for 'Best' (minimum variance) when:

A. There are many variables

✓ **B. Homoskedasticity is violated (Heteroskedasticity present)**

C. Sample size is large

D. The model is log-linear

Lecture 15: Multicollinearity

Q479. Under perfect multicollinearity, the normal equations are:

A. Independently solvable

✓ **B. Linearly dependent — cannot be uniquely solved for all parameters**

C. Zero

D. Quadratic

Lecture 6: Simple Regression Model & OLS

Q480. For a consistent estimator, as n → ∞:

A. The estimator becomes more biased

✓ **B. The estimator converges in probability to the true parameter**

C. The variance increases

D. R^2 decreases

Lecture 7: Estimation, Testing & R^2 in Regression

Q481. A 95% confidence interval for β1 means:

A. There is 95% probability β1 is in this interval for this sample

✓ **B. In repeated sampling, 95% of such intervals would contain the true β1**

C. β1 is significant at 5% level

D. R^2 is at least 95%

Lecture 18: Heteroskedasticity: Formal Tests

Q482. The White test is run using the auxiliary regression $e^2 = \gamma 0 + \gamma 1 \blacksquare + \gamma 2 \blacksquare^2 + v$ as a simplified alternative because:

A. It is always more powerful than the full auxiliary regression

✓ **B. It avoids having too many regressors when K is large — uses fitted values as a compact proxy**

C. It does not need the original X variables

D. It tests for multicollinearity instead

Lecture 5: Multivariate Optimization & Probability Review

Q483. For the profit function $\pi = 80x - 2x^2 - xy - 3y^2 + 100y$ with critical values $x\blacksquare=16.52$, $\blacksquare=13.91$, and $\pi xy = -1$, the Hessian condition $\pi xx \cdot \pi yy > \pi xy^2$ gives:

✓ **A. $(-4)(-6) = 24 > (-1)^2 = 1$, so condition is satisfied**

B. $(-4)(-6) = 24 < 1$, so condition fails

C. $(4)(6) = 24 = 1^2 = 1$, borderline

D. Cannot be determined without more information

Lecture 11: Transformation & Functional Forms in Regression

Q484. In the linear-in-variable model Y = a + bX, both marginal effect (b) and elasticity (bX/Y) change with X because:

A. b changes at every point
✓ **B. The marginal effect is constant (b) but elasticity (bX/Y) varies with X and Y values**
C. Both are constant in linear models
D. Only elasticity is constant in linear models

Lecture 7: Estimation, Testing & R^2 in Regression

Q485. The LINEST formula in Excel returns results as an ARRAY. This means:

A. Only one value is returned
✓ **B. Multiple statistics are returned in a range of cells — you must select the array and press Ctrl+Shift+Enter**
C. Results are displayed as a chart
D. The formula requires the Data Analysis ToolPak to be installed first

Lecture 5: Multivariate Optimization & Probability Review

Q486. For a Chi-square variable with k = 4 degrees of freedom, the mean equals:

A. 2
✓ **B. 4**
C. 8
D. 16

Lecture 18: Heteroskedasticity: Formal Tests

Q487. The Harvey-Godfrey test differs from the Park test in its auxiliary regression in that it uses:

A. lnX as the regressor
✓ **B. X (the original variable) instead of lnX as the regressor**
C. The absolute residual instead of squared residual
D. Y instead of X as the regressor

Lecture 6: Simple Regression Model & OLS

Q488. OLS requires 'Sample Variation in Explanatory Variable Xi'. This means:

A. All Xi must be the same
✓ **B. Xi must not be constant — there must be variation in the observed X values**
C. Xi must follow a normal distribution
D. Xi must be positive

Lecture 15: Multicollinearity

Q489. In the Pakistan GDP regression, including both exports and imports jointly caused multicollinearity because:

A. They are the same variable
✓ **B. They both share a common upward trend over time and are correlated with each other**
C. They were measured in different currencies
D. They had missing values

Q490. The Tolerance = 1 – R^2 (from auxiliary regression) close to ZERO means:

A. The variable is perfectly independent of others

✓ **B. The variable is almost entirely explained by other regressors — severe MC**

C. There is no multicollinearity

D. R^2 of the main regression is zero

Lecture 9: Multiple Regression

Q491. A researcher estimates Y = 264.67 − 0.037X1 − 5.17X2. What does the negative sign on X2 imply?

A. X2 has no relationship with Y

✓ **B. Holding X1 constant, a one-unit increase in X2 is associated with a 5.17 unit DECREASE in Y on average**

C. X2 should be dropped from the model

D. The model is incorrectly specified

Lecture 7: Estimation, Testing & R^2 in Regression

Q492. The sum of squared residuals $\sum e_i^2 = 0$ would imply:

A. The model explains none of the variation

✓ **B. All observations lie exactly ON the regression line (perfect fit, $R^2=1$)**

C. The intercept is zero

D. The slope is zero

Lecture 10: K-Variable Regression & Analysis ToolPak

Q493. In regression analysis, increasing the number of independent variables from 2 to 3 while keeping N constant will:

A. Increase degrees of freedom by 1

✓ **B. Decrease degrees of freedom by 1**

C. Not affect degrees of freedom

D. Double the degrees of freedom

Lecture 5: Multivariate Optimization & Probability Review

Q494. The condition for a 'Saddle Point' in multivariate optimisation is that the second order partials:

A. Are both positive

B. Are both negative

✓ **C. Have opposite signs (one positive, one negative)**

D. Are both equal to zero

Lecture 6: Simple Regression Model & OLS

Q495. Which property of estimators is NOT part of 'BLUE'?

A. Best (minimum variance)

B. Linear

C. Unbiased

✓ **D. Exact (zero estimation error)**

Lecture 11: Transformation & Functional Forms in Regression

Q496. In the log-level (semi-log) model $\ln Y = \beta 0 + \beta 1 X$, the slope $\beta 1$ approximately gives the:

A. Absolute change in Y per unit change in X

✓ **B. Proportionate (percentage/100) change in Y per unit change in X**

C. Elasticity of Y with respect to X

D. Level of Y when X = 0

Lecture 14: Logarithms in Regression & Missing Data

Q497. The method of 'Pairwise Deletion' is also called:

A. Complete case analysis

✓ **B. Available case analysis**

C. Multiple imputation

D. Single imputation

Lecture 6: Simple Regression Model & OLS

Q498. In OLS regression, the estimate b is a 'LINEAR estimator' because:

A. The model equation is linear in X

✓ **B. The OLS slope b can be written as a linear combination of the observed Y values**

C. The intercept is constant

D. The error term is zero

Lecture 13: Qualitative Variables & Dummy Variables

Q499. A regression model with dummy variables for EACH category AND an intercept will have perfect multicollinearity. This is called:

A. Omitted variable bias

✓ **B. The Dummy Variable Trap**

C. Heteroskedasticity

D. The Goldfeld-Quandt trap

Lecture 18: Heteroskedasticity: Formal Tests

Q500. The chi-square goodness-of-fit statistic LM = N × R^2 from the Breusch-Pagan test is compared to:

A. t-table with n–k degrees of freedom

✓ **B. Chi-square table with K–1 degrees of freedom (K = number of X variables in auxiliary regression)**

C. F-table with k–1, n–k degrees of freedom

D. Standard normal z-table