Name: Enrolment No:	

University of Petroleum & Energy Studies (UPES)
School of Business (SoB)
End-Semester Examination – December 2023

Program: BBA Foreign Trade | **Semester: III**
Subject / Course: Introduction to Econometrics | **Maximum Marks: 100**
Course Code: ECON 2037 | **Duration: 03 Hours**

INSTRUCTIONS:

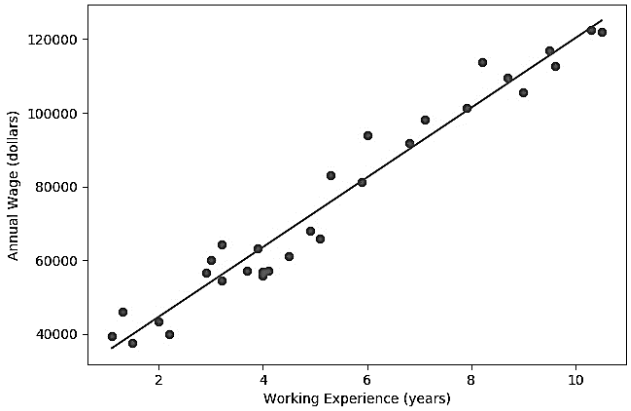
- **This is a CLOSED-BOOK EXAM. Only Non-scientific calculator is allowed.**
- **Cellphones / Tablets / Laptops / Books / Notes etc. are NOT allowed.**
- **All questions are compulsory. If Choice is there, it is *indicated within the question as* OR.**
- **Your answers must be "brief & to the point."**

Q. No.	Questions	Marks	COs
	SECTION A **10Q x 2M = 20 Marks**		
Q 1.	Two events, A and B, are said to be mutually exclusive if: **A)** $P(A \mid B) = 1$ **B)** $P(B \mid A) = 1$ **C)** $P(A \,\&\, B) = 1$ **D)** $P(A \,\&\, B) = 0$	2	CO1
Q 2.	Type I error occurs when we: **A)** reject a false null hypothesis. **B)** reject a true null hypothesis. **C)** do not reject a false null hypothesis. **D)** do not reject a true null hypothesis.	2	CO1

Q 3.	The violation of the assumption of constant variance of the residual is known as: **A)** The variance of the errors is not constant. **B)** The variance of the dependent variable is not constant. **C)** The errors are not linearly independent of one another. **D)** The errors have non- zero mean.	2	CO1
Q 4.	Autocorrelation is generally occurred in: **A)** Cross-section data. **B)** Time series data. **C)** Pooled data. **D)** None of the above.	2	CO1
Q 5.	In the regression function $\boxed{Y = \alpha + \beta X + \varepsilon}$: **A)** X is the regressor. **B)** Y is the regressor. **C)** α is the regressor. **D)** ε is the regressor.	2	CO1
Q 6.	BLUE is referred as the **A)** Best Linear Unbiased Estimator. **B)** Best Linear Unconditional Estimator. **C)** Basic Linear Unconditional Estimator. **D)** Both **B)** and **C)**.	2	CO1
Q 7.	Data on one/more variables collected at a given point of time is known as: **A)** Panel data. **B)** Time series data. **C)** Pooled data. **D)** Cross-section data.	2	CO1
Please Turn Over			

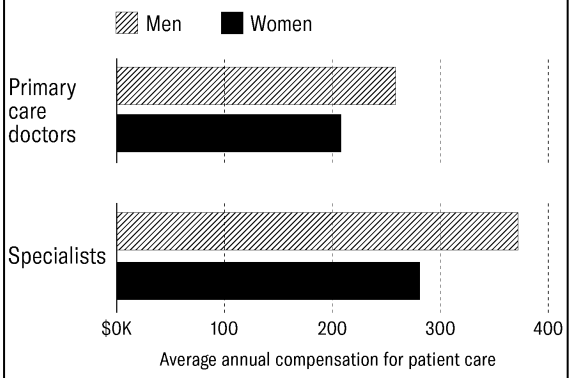
Q 8.	Probability of occurrence of an event lies between **A)** –1 and 0. **B)** –1 and 1. **C)** 1 and 0. **D)** –100 and 100.	2	CO1
Q 9.	A sure way of removing multicollinearity from the model is to: **A)** Work with panel data. **B)** Drop variables that cause multicollinearity in the first place. **C)** Transform the variables by first order of differencing them. **D)** Obtaining additional sample data.	2	CO1
Q 10.	The coefficient of determination, R^2 shows: **A)** The proportion of the variation in the dependent variable Y is explained by the independent variable X. **B)** The proportion of the variation in the dependent variable X is explained by the independent variable Y. **C)** The proportion of the variation in ε is explained by the independent variable X. **D)** Both **A)** and **C)**.	2	CO1
SECTION B **4Q x 5M = 20 Marks**			
Q 11.	True or False? Briefly justify the reasoning. *"If a fair coin is tossed many times for independent trials, and the last eight tosses are all tails, then the chance that the next toss will be tails is somewhat less than 50%."*	5	CO2
Q 12.	What is a *Null Hypothesis* (H_0) and an *Alternative Hypothesis* (H_1)? Using a relevant example, briefly explain these two concepts.	5	CO2
Please Turn Over			

Q 13.	Using a relevant example, briefly explain the difference between *Two-Tailed* & *One-Tailed* Tests.	5	CO2
Q 14.	A recent research survey done by Dr. Chakraborty asked 15,292 randomly sampled registered Indian voters about their political affiliation (Rightist, Leftist, or Independent) and whether or not they identify as 'Swing Voters.' - 15% of respondents identified as Independent, - 33% identified as Swing Voters, and - 21% identified as both. What percent of voters are Independent <u>OR</u> Swing Voters? Show your calculation.	5	CO2
	SECTION C **3Q x 10M = 30 Marks**		
Q 15.	Below are the final exam scores of 20 *Introductory Econometrics* students. 79 83 57 82 94 83 72 74 73 71 66 89 78 81 78 81 88 69 77 79 **1.1.**What is the mean score? [<u>2 points</u>] **1.2.**What is the median score? [<u>2 points</u>] **1.3.**What is the mode? [<u>2 points</u>] **1.4.**What is the Standard Deviation (S.D.)? [<u>2 points</u>] **1.5.**Draw a free-hand histogram for the score distribution. [<u>2 points</u>] **OR** Using the regression function $\boxed{Y_i = \alpha + \beta X_i + \varepsilon_i}$ write down the **key five assumptions of the Ordinary Least Squares (OLS)** and briefly explain each one of them.	10	CO3
	Please Turn Over		

Q 16.	Briefly explain the following concepts: **16.1.** Leptokurtic distribution **16.2.** Covariance **16.3.** Mean squared prediction error (MSPE) **16.4.** i.i.d. random variables **16.5.** The 'central limit theorem'	10	CO3
Q 17.	To study the relationship between the wage (dependent variable) and working experience (independent variable), we use a linear regression model: $\boxed{Wage_i = \alpha + \beta * Experience_i + \varepsilon_i}$. In this study, we use 30 data points [$i = 1 \rightarrow 30$], where the annual salary (in USD) ranges from \$39,343 to \$1,21,872 and the years of experience range from 1.1 years to 10.5 years. Looking at this scatterplot below, we can imagine that the relationship in this sample is pretty close to linear.  A quick MS Excel regression exercise spits out the following equation: $\boxed{\widehat{\boldsymbol{Wage}}_{\boldsymbol{i}} = 25792.20 + 9449.96 * Experience_i}$. **Calculate and interpret** the values of $\boxed{\widehat{\boldsymbol{Wage}}_{\boldsymbol{i}}}$ for a typical worker i from the sample: **1.** When $Experience_i$ $\boxed{\uparrow}$ by 1 extra year, i.e., $\frac{\partial Y_i}{\partial X_i}$. [1 point] **2.** When $Experience_i = 0$ year. [3 points] **3.** When $Experience_i = 5$ years. [3 points] **4.** When $Experience_i = 10$ years. [3 points]	10	CO3
	Please Turn Over		

SECTION D
2Q x 15M = 30 Marks

Q 18.	The cost of attending your university has once again gone up. Although you have been told that education is investment in human capital, which carries a return of roughly 10% a year, you (and your parents) are not pleased. One of the administrators at your university does not make the situation better by telling you that you pay more because the reputation of your institution is better than that of others. To investigate this hypothesis, you collect data randomly for 100 Indian universities and business schools from the 2023 NIRF ranked institutions. Next you perform a linear regression analysis on MS Excel, that you just learned in the *Introduction to Econometrics* class. You find the following regression equation: $\widehat{Cost} = 7{,}311.17 + 35{,}985.20 \times Reputation - 5{,}508.20 \times Student_{Size} + 98{,}406.79 \times Private$ where, $Cost = Tution + Fees + Hostel\ charges + Food\ expenses$ $Reputation = An\ index\ (valued\ 1-5)\ based\ on\ academic\ reputation\ in\ faculty\ \&\ placements$ $Student_{Size} = Total\ number\ of\ Business\ students\ (BBA + MBA)$ $Private = Whether\ the\ university\ is\ private\ OR\ public/government\ institution?$ **(a)** Identify what type of variables each of the above is. **(b)** Interpret the results using the numbers (the coefficients) in the above equation. **(b.1.)** $+ 7{,}311.17$ → **(b.2.)** $+ 35{,}985.20$ → **(b.3.)** $- 5{,}508.20$ → **(b.4.)** $+ 98{,}406.79$ → **(c)** Now, you remove the $Student_{Size}$ variable from MS Excel regression and find a new equation: $\widehat{Cost} = 5{,}450.35 + 35{,}538.84 \times Reputation + 5{,}50{,}935.70 \times Private$ → Why do you think that the effect of attending a private university has suddenly increased from $Rs.\,98{,}406.79$ to $Rs.\,5{,}50{,}935.70$ now? Briefly explain.	15	CO4

Q 19.	While we celebrate the *"Happy"* International Women's Day every 8th of March with *archies cards, discounts on spa & beauty products, glorifying femininity & motherhood*, we probably forget the history of struggle against all of these. Carefully examine this figure (on the right-hand). Now, propose a typical regression model for the research question: *"whether gender matters for compensation/salary in the same profession?"* in the format of $\boxed{Y_i = \beta_0 + \beta_1 X_i + u_i}$; where $\boxed{Y_i}$ is the value of 'dependent' variable for $\boxed{i^{th}}$ person, $\boxed{\beta_0}$ is the constant/intercept, $\boxed{\beta_1}$ is the slope/coefficient of $\boxed{X_{ij} = X_{i1}, X_{i2}, X_{i3}, etc.}$ the 'independent' variables, and $\boxed{u_i}$ is the error term of the regression. **Doctors' Glaring Pay Gap** Across the board, women physicians in the U.S. make substantially less than their male counterparts.  Source: Medscape Physician Compensation Report 2019 HBR *Note: Please mention what exactly your variables are? What type of variables are they, i.e., binary, categorical, continuous, etc.? And, what values they could be assigned with?* *Hint:* $\boxed{Y_i}$ = *"average annual compensation (in $)"* ⟹ *a continuous variable* ⟹ *values ranging from $0K to $400K.*	15	CO4