



**NAVRACHANA  
UNIVERSITY**

a UGC recognized University

**School:** School of Engineering and Technology  
**Program/s:** BSc ( Data Science )  
**Year:** 2<sup>nd</sup> **Semester:** III  
**Examination:** End Semester Examination  
**Examination year:** December - 2021

**Course Code:** DS207 **Course Name:** Econometrics  
**Date:** 9/12/2021  
**Time:** 8:30 am to 10:30 am

**Total Marks:** 40  
**Total Pages:** 2

**Instructions:**

- Write each answer on a new page.
- Use of a calculator is permitted.
- Draw all relevant waveforms in answer sheet only.
- \* COs=Course Outcome mapping. # BTL=Bloom's Taxonomy Level mapping

Q. No.	Details	Marks	COs*	BTL#																											
Q.1(a)	<p><b>Answer in Short:</b></p> <p>(1) In simple linear regression model, <math>y = b_0 + b_1 x + u</math>, what does <math>u</math> represent?</p> <p>(2) In a model, <math>wage = b_0 + b_1 educ + u</math>, what does <math>b_1</math> measure? Where wage in dollars per hour and educ in years of education.</p> <p>(3) What is interpretation of <math>E(u) = 0</math>?</p> <p>(4) State the relation among SST, SSE and SSR.</p>	4	CO1	BT1, BT2																											
Q.1(b)	<p><b>Attempt Any TWO:</b></p> <p>(i) Explain the concepts: (a) Linearity in parameters (b) Linearity in variables  (c) Population Regression Function (d) Sample Regression Function</p> <p>(ii) Explain the method of ordinary least squares to obtain the best values of parameters in sample regression model and derive the expression for the best values of parameters.</p> <p>(iii) The following table contains the ACT scores and the GPA (grade point average) for eight college students. Grade point average is based on a four-point scale and has been rounded to one digit after the decimal.</p> <table border="1"> <thead> <tr> <th>Student</th> <th>GPA</th> <th>ACT</th> </tr> </thead> <tbody> <tr><td>1</td><td>2.8</td><td>21</td></tr> <tr><td>2</td><td>3.4</td><td>24</td></tr> <tr><td>3</td><td>3.0</td><td>26</td></tr> <tr><td>4</td><td>3.5</td><td>27</td></tr> <tr><td>5</td><td>3.6</td><td>29</td></tr> <tr><td>6</td><td>3.0</td><td>25</td></tr> <tr><td>7</td><td>2.7</td><td>25</td></tr> <tr><td>8</td><td>3.7</td><td>30</td></tr> </tbody> </table> <p>Estimate the relationship between GPA and ACT using OLS; that is, obtain the intercept and slope estimates in the equation <math>GPA = \hat{b}_0 + \hat{b}_1 ACT</math>. Comment on</p>	Student	GPA	ACT	1	2.8	21	2	3.4	24	3	3.0	26	4	3.5	27	5	3.6	29	6	3.0	25	7	2.7	25	8	3.7	30	16	CO1, CO2, CO3	BT1, BT2, BT5
Student	GPA	ACT																													
1	2.8	21																													
2	3.4	24																													
3	3.0	26																													
4	3.5	27																													
5	3.6	29																													
6	3.0	25																													
7	2.7	25																													
8	3.7	30																													

	the direction of the relationship. How much higher is the GPA predicted to be if the ACT score is increased by five points?			
<b>Q.2(a)</b>	<p><b>Answer in short:</b></p> <p>(1) Write formula for goodness of fit.</p> <p>(2) State the relation between coefficient of correlation and goodness of fit.</p> <p>(3) For Sample Linear Regression Model, <math>y = -0.912 + 2.25x</math>, <math>r^2 = 0.440</math>. How do you interpret <math>r^2</math> ?</p> <p>(4) What is meaning of homoscedasticity assumption, <math>\text{Var}(u/x) = \sigma^2</math>?</p> <p>(5) For which purpose we construct confidence intervals in regression analysis?</p>	<b>5</b>	CO2	BT1, BT2
<b>Q.2(b)</b>	<p><b>Attempt ANY THREE:</b></p> <p>(i) State the formulas for standard errors and variance of sample parameters <math>b_1</math>, <math>b_2</math>.</p> <p>(ii) Consider the sum: <math>\sum yx = 131.7856</math>, <math>\sum x^2 = 182.0</math>, <math>\sum u^2 = 9.83017</math>, <math>n = 13</math>, <math>\bar{x} = 0.7240967</math>, <math>\bar{y} = 8.674708</math>. Calculate sample parameters <math>b_1</math>, <math>b_2</math>, sample variance, variance and standard error of parameter <math>b_2</math>.</p> <p>(iii) The following model is a simplified version of the multiple regression model used by Biddle and Hamermesh (1990) to study the tradeoff between time spent sleeping and working and to look at other factors affecting sleep:</p> <p style="padding-left: 40px;"><math>\text{sleep} = b_0 + b_1 \text{totwrk} + b_2 \text{educ} + b_3 \text{age} + u</math>,</p> <p>where sleep and totwrk (total work) are measured in minutes per week and educ and age are measured in years.</p> <p style="padding-left: 40px;">(1) What signs do you think <math>b_2</math> and <math>b_3</math> will have?</p> <p style="padding-left: 40px;">(2) Would you say totwrk, educ, and age explain much of the variation in sleep? What other factors might affect the time spent sleeping? Are these likely to be correlated with totwrk?</p> <p>(iv) State the assumptions required for unbiasedness of OLS estimators. State mathematical form of unbiasedness of OLS.</p>	<b>15</b>	CO3, CO4	BT2, BT3, BT5

\*\*\*\*\*End of Question Paper\*\*\*\*\*