

2021

STATISTICS — HONOURS — PRACTICAL

Paper : DSE-A-1P

(Econometrics)

Full Marks : 30

*The figures in the margin indicate full marks.*

1. In this exercise a researcher uses data on NBA players' salaries and their determinants. She is interested in knowing the effect of performance on NBA players' salaries. The following information is available for 56 NBA players.

Table 1 : Variables of the dataset— names and description

SALARY	Salary earned by players in thousands of dollars.
HT	Height of the players in inches.
WT	Weight of each player in pounds.
AGE	Age of each player.
MIN	Number of minutes that each player played during the season.
STEALS	Number of times that player stole ball from opponents.
BLOCKS	Number of blocked shots.
POINTS	Number of points that the player scored in the full season.

The summary statistics for all the variables in Table 1, as well as the correlation matrix, are presented in

Table 2 : Summary statistics

Variable	Mean	Std. Dev.	Minimum	Maximum
SALARY	1668.04	667.910	1000	3750
HT	80.6250	3.66091	73	88
WT	226.804	26.7034	175	290
AGE	28.4107	3.03181	23	36
MIN	2538.96	670.669	189	3255
STEALS	97.9821	85.9193	6	564
BLOCKS	70.5179	74.0361	5	315
POINTS	1369.68	578.186	116	2633

Please Turn Over

Table 3 : Correlation matrix

	SALARY	HT	WT	AGE	MIN	STEALS	BLOCKS	POINTS
SALARY	1	0.003	0.048	-0.075	0.094	0.082	0.088	0.233
HT		1	0.832	0.2926	-0.345	-0.349	0.556	-0.365
WT			1	0.086	-0.184	-0.225	0.473	-0.289
AGE				1	-0.112	-0.346	0.169	-0.081
MIN					1	0.319	0.048	0.793
STEALS						1	-0.128	0.38
BLOCKS							1	-0.078
POINTS								1

Answer the following questions :

- Is this a cross-section or a time series? Why? What is the unit of analysis in this dataset?
- How old is the youngest basketball player of this sample?
- Could you tell which variable is more dispersed by looking at the values of the standard deviations in Table 2 ?
- Could you say, by looking at Table 3, that there is a penalty in terms of lower wages associated with age? Explain.
- Which variables have the highest correlation (positive or negative) with wage? Explain. 5

2. A company A operates with the following production function :

$$Y_t^A = 110 + 0.65K_t^A \left( R_A^2 = 0.37 \right)$$

such that  $Y_t^A$  measures total production in thousand Euros in year  $t$  and  $K_t^A$  measures the use of capital in thousand Euros in year  $t$ .

- Interpret the coefficients of the estimated production function.
- A competitor of company A, company B operates according to a different production function defined as :

$$Y_t^B = 80 + 0.50K_t^B \left( R_B^2 = 0.48 \right)$$

Interpret the coefficients of the estimated production function for company B in comparison to the coefficients for company A.

- (c) In 2010 ( $t = 2010$ ), the use of capital in company A had a value of 320,000 Euros and 280,000 Euros in company B. Both companies are planning to expand their businesses to the Brazilian market in 2015. Therefore, their capital levels will increase 20% respect to 2010. Find the total production prediction in 2015 ( $t = 2015$ ) for each company using the estimated cost functions. Explain which company will obtain more accurate prediction in your opinion.
- (d) Do you think the relationship between production and the use of capital has constant returns (whether linearity assumption is satisfied)? If no, specify a more realistic regression model.

2+2+2+2

3. Using the data of eight firms, a regression model was estimated to analyze the relationship between investment in thousand Euros ( $y_t$ ) and production growth rate  $\%(x)$  :

$$\hat{y}_t = 3.841 - 0.0812x_t, R^2 = 0.466, SSR = 39.21, n = 8$$

s.e (2.12) (0.038)

Additionally, two different regressions are estimated. The first one only takes into account European firms within the original sample :

$$\hat{y}_t = -0.372 + 0.108x_t, R^2 = 0.976, SSR = 0.949, n = 4$$

s.e (0.782) (0.012)

and the second one only takes into American firms within the original sample :

$$\hat{y}_t = 1.259 + 0.171x_t, R^2 = 0.933, SSR = 1.407, n = 4$$

s.e (1.43) (0.032)

Find whether making the distinction between European and American firms helps to understand better the behaviour of investment and *interpret* your results. How you will include the distinction (if it is better to consider) in a single regression equation. 3+3

4. The following table shows two different samples with two explanatory variables each of them in order to study the behaviour of Y (dependent variable) :

**Sample 1    Sample 2**

Observation	Y	X1	X2	Z1	Z2
1	1	2	4	2	4
2	4	6	12	6	12
3	2	4	11	4	8

- (a) Can you *detect* a multicollinearity problem in any of the two samples?
- (b) If yes, please *explain* the consequences in your OLS estimations in each sample.
- (c) If yes, please *explain* the strategies that you would use in order to solve the problem in each sample. 3+4+4