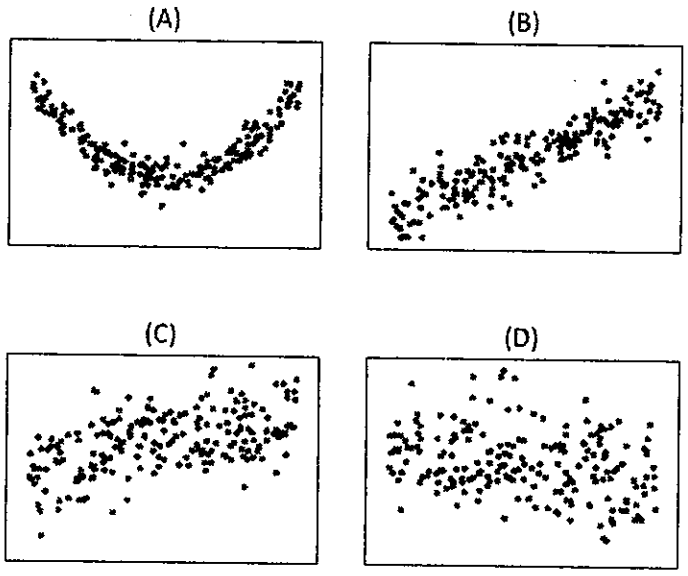
※ 注意：請於試卷內之「非選擇題作答區」依序作答，並應註明作答之大題及小題題號。

1. **(5 points)** Match the calculated correlations to the corresponding scatterplot. Note that the scatterplots have the same scale.

Correlation Coefficient	Scatterplot
0.5285	?
-0.2502	?
0.9177	?
-0.0150	?



2. Let X and Y have joint probability density function (pdf):

$$f(x,y) = \begin{cases} cy & 0 < y < x < 2 \\ 0 & \text{elsewhere} \end{cases}$$

(a) **(6 points)** Find the value of c that makes $f(x,y)$ a valid pdf.

(b) **(6 points)** Find $E(Y \mid X = 5)$.

3. In each part below, there is a value of interest and two scenarios (I and II). For each part, determine if the value of interest is larger under scenario I, scenario II, or whether the value is equal under the scenarios. Explain your answers in English. (Answers in other languages will NOT be graded.)

(a) **(5 points)** The standard error of a sample mean when the sample standard deviation is 120 and the number of observations is (I) 50 or (II) 100.

(b) **(5 points)** The margin of error of a confidence interval when the confidence level is (I) 95% or (II) 90%.

(c) **(5 points)** The power of a hypothesis test when the significance level is (I) 5% or (II) 10%.

4. Emily collects data on the monthly salaries of 300 people in Taipei. She makes a table for the monthly salaries in thousands of New Taiwan Dollars (NTD) by occupation.

Occupation	Number of obs	Mean	Std. Dev.
Doctor	100	150	37
Engineer	100	105	26
Lawyer	100	120	41

見背面

(a) **(3 points)** Let S_i denote individual i's monthly salary in thousands of NTD. Use the data to estimate regression model:

$$S_i = \alpha + \varepsilon_i.$$

What is the least squares estimate of $\hat{\alpha}$?

(b) **(6 points)** Let ENG_i be a binary variable equal to 1 if individual i is an engineer and 0 otherwise. Use the data to estimate regression model:

$$S_i = \alpha + \beta ENG_i + \varepsilon_i.$$

What is the least squares estimate of $\hat{\alpha}$ and $\hat{\beta}$?

(c) **(9 points)** Let LAW_i be a binary variable equal to 1 if individual i is a lawyer and 0 otherwise. Use the data to estimate regression model:

$$S_i = \alpha + \beta ENG_i + \gamma LAW_i + \varepsilon_i.$$

What is the least squares estimate of $\hat{\alpha}$, $\hat{\beta}$, and $\hat{\gamma}$?

5. **(10 points)** Assume that the true model is

$$y = \beta_1 x_1 + \beta_2 x_2 + u,$$

where $E(x_1 x_2) \neq 0$, $E(x_1 u) = 0$, and $E(x_2 u) = 0$. If we regress y on x_1 and x_2 (without a constant term) separately to obtain $\hat{\beta}_1$ and $\hat{\beta}_2$. Are these two estimators consistent? Show your answer in details.

6. To study airfares in the United States, the data include 1149 domestic routes from 1997 to 2000 (annual data). Consider the following model:

$$fare_{it} = \beta_0 + \beta_1 dist_i + \beta_2 concen_{it} + \beta_3 y98_t + \beta_4 y99_t + \beta_5 y00_t + u_{it},$$
$$i = 1, 2, ..., 1149;\ t = 1997, 1998, 1999, 2000$$

where $fare_{it}$ is the average one-way fare (US dollars), $dist_{it}$ is the distance for the route (in miles), and $concen_{it}$ is the market share for the biggest carrier on that route. $y98_t$, $y99_t$, $y00_t$ are year dummies for 1998, 1999, and 2000. Table 1 shows the summary statistics for the data.

Table 1: Summary Statistics

Variables	Obs	Mean	Std. Dev.	Min	Max
fare	4,596	178.8	74.88	37	522
dist	4,596	989.7	611.8	95	2,724
concen	4,596	0.610	0.196	0.160	1
y98	4,596	0.250	0.433	0	1
y99	4,596	0.250	0.433	0	1
y00	4,596	0.250	0.433	0	1

(a) **(6 points)** From column (2) in Table 2, please interpret the coefficients on *dist* and *concen.*

(b) **(6 points)** For the coefficients on *dist*, please explain the difference between column (1) and column (2) in Table 2.

(c) **(6 points)** If someone claims that there is no change on airfares over time, how can you specify a null hypothesis testing? Do the hypothesis testing based on column (2) and column (3) in Table 2, and conclude your results.

接次頁

Table 2: Regression Results

Variables	(1) *fare*	(2) *fare*	(3) *fare*	(4) *fare*	(5) log(*fare*)
dist	0.0763***	0.0888***	0.0890***	0.0823***	
	[0.00141]	[0.00163]	[0.00162]	[0.00559]	
concen		73.36***	74.64***	73.79***	0.324***
		[5.065]	[5.044]	[5.089]	[0.0302]
y98			1.023	1.030	0.0214
			[2.377]	[2.377]	[0.0142]
y99			4.791**	4.785**	0.0376***
			[2.377]	[2.377]	[0.0142]
y00			15.09***	15.08***	0.0995***
			[2.377]	[2.377]	[0.0142]
log(*dist*)					0.454***
					[0.00900]
$dist^2$				2.60e-06	
				[2.07e-06]	
Constant	103.3***	46.19***	39.97***	43.61***	1.819***
	[1.643]	[4.255]	[4.491]	[5.346]	[0.0726]
Observations	4,596	4,596	4,596	4,596	4,596
R-squared	0.389	0.416	0.422	0.422	0.392

Standard errors in brackets, *** p<0.01, ** p<0.05, * p<0.1

(d) **(6 points)** If we also include the square term for *dist* in column (4) in Table 2, please compute the average marginal effect of *dist* on *fare*.

(e) **(6 points)** If we take the log both on *fare* and *dist* in column (5) in Table 2, how can you interpret the coefficient on log(*dist*)?

(f) **(10 points)** Assume that a new airline entered into the market in 1999, and only some of routes are affected by this entry. If we want to know whether this entry increased the competition and lowered down the average airfares, could you construct a new model to verify this effect? Please write down the model and indicate which coefficients can help you conclude your results. You can define any variables yourself.

試題隨卷繳回