

※本大題請於試卷內之「選擇題作答區」依序作答。

1. [30分，選擇題，共15小題]

「A random variable X is classified as five levels, 1-5. One would like to __ (1) __ the five levels of this random variable into a new binary variable, say Z. There is another random variable called Y, also a binary variable, which is the outcome of interest in this study. One is intended to investigate the __ (2) __ between Z and Y using a statistical method [__ (3) __]. The p-value less than 5% is set as a statistically __ (4) __ level to test this __ (5) __. After such a __ (6) __ analysis, another principal investigator is worried about whether another independent variable C (also a binary variable) may __ (7) __ the analysis between Z and Y mentioned above. To __ (8) __ for the variable C, the __ (9) __ is considered with the __ (10) __ of __ (11) __ into the model besides the independent variable of main interest. The main effect can be estimated by taking the __ (12) __ of regression coefficient, reflecting the __ (13) __ of the relationship between Z and Y. If the main effect is also dependent on another independent variable, then the __ (14) __ term needs to be included in the model. The model is therefore called the __ (15) __ model.」
根據上述短文回答下列問題。

[(1)~(15) 為選擇題，依上文之標示題號依序作答，每題 2 分。]

- (1) (A) divert, (B) categorize, (C) collapse, (D) none of above.
- (2) (A) connotation, (B) connection, (C) association, (D) refaction.
- (3) which one is not appropriate : (A) Chi-square method, (B) two-sample proportion test, (C) a simple linear regression, (D) logistic regression.
- (4) (A) associated, (B) relevant, (C) saturated, (D) non of above.
- (5) (A) postulate, (B) clue, (C) annexation, (D) construct.
- (6) (A) simple, (B) crude, (C) net, (D) delicate.
- (7) (A) relate, (B) simplify, (C) confound, (D) balance.
- (8) (A) test, (B) swap, (C) settle, (D) none of above.
- (9) (A) logistic regression model, (B) Cox regression model, (C) linear regression model, (D) none of above.
- (10) (A) deletion, (B) amalgamation, (C) incorporation, (D) none of above.
- (11) (A) X, (B) Z, (C) C, (D) Y.
- (12) (A) logarithm, (B) differential, (C) exponential, (D) transformation.
- (13) (A) attenuation, (B) strength, (C) specification, (D) distraction.
- (14) (A) regression, (B) interaction, (C) exchange, (D) distortion.
- (15) (A) parsimonious, (B) reduced, (C) modified, (D) saturated.

見背面

2. [35分，共5小題，每題7分]

以英文解釋以下名詞。

- (1) 「sample space」
- (2) 「random variable」
- (3) 「random sample」
- (4) 「parameter」
- (5) 「statistical inference」

3. [35分]「對一組配對資料 $(y_i, x_i), i=1, 2, \dots, n$ ，將 y_i 視為依變數的 Y 的一個觀察值， x_i 視為自變數 X 的一個觀察值， Y 和 X 之間可以線性模式描述：

觀察值， x_i 視為自變數 X 的一個觀察值， Y 和 X 之間可以線性模式描述：

$$y_i = \alpha + \beta x_i + \varepsilon_i$$

其中， α 是模式的截距參數； β 是斜率參數，又稱為回歸係數； ε_i 是隨機誤差。估計 α 和 β 可以最大概似法或最小平方方法進行，所得估計式可以證明具有最佳線性無偏性質。簡單線性回歸模式可以推廣至複回歸模式，亦即模式中放入多於一個的自變數。當自變數有多個時，為達到模式精簡的要求，必須進行模式選取過程，透過一些選取標準，選取出一個適合的模式。回歸模式的發展起源來自於遺傳資料的分析，現在則被廣泛地應用於社會學、工程學、醫學等各個領域。」

將上述文字譯成英文。

試題隨卷繳回