

科目：專業英語

系所組：統計資訊學系應用統計碩士班

I. Chinese-English Translation (中翻英)

1. 平均數的計算為觀測值的加總，再除以觀測值的個數。(15%)
2. 假設大學生每晚平均睡眠為 7.1 小時，且標準差為 35 分鐘。如果睡眠時間為常態分配，則大學生睡眠低於 7 小時的比例為何？(15%)

II. English-Chinese Translation (英翻中)

(A).

1. MS EXCEL provides statistical functions to compute descriptive statistics for a data set. These functions can be used to compute one statistic at a time (e.g., mean, variance, etc.). (15%)
2. Correlation coefficient is a measure of linear association between two variables that takes on values between -1 and $+1$. Values near $+1$ indicate a strong positive linear relationship; values near -1 indicate a strong negative linear relationship; and values near zero indicate the lack of a linear relationship. (15%)
3. Standard deviation is a measure of variability computed by taking the positive square root of the variance. (15%)

(B). Choose one of the following questions to answer. (任選一題 [註：4(A) 或 4(B)] 作答) (25%)

4(A). Journal article 1. Data learning from big data (Source: *Statistics & Probability Letters*, 2018, 136, pp. 15-19.)

(英翻中)

Technology is generating a huge and growing availability of observations of diverse nature. This big data is placing data learning as a central scientific discipline. It includes collection, storage, preprocessing, visualization and, essentially, statistical analysis of enormous batches of data. In this paper, we discuss the role of statistics regarding some of the issues raised by big data in this new paradigm and also propose the name of data learning to describe all the activities that allow to obtain relevant knowledge from this new source of information.

4(B). Journal article 2. Forecasting daily Covid-19 cases in the world with a hybrid ARIMA and neural network model (Source: *Applied Soft Computing*, 2022, 126, 109315.)

(英翻中)

The use of models to predict disease cases is common in epidemiology and related areas, in the context of Covid-19, both ARIMA and Neural Network models can be applied for purposes of optimized resource management, so the aim of this study is to capture the linear and non-linear structures of daily Covid-19 cases in the world by using a hybrid forecasting model. In summary, the proposed hybrid system methodology consists of two steps. In the first step, an ARIMA model is used to analyze the linear part of the problem. In the second step, a neural network model is developed to model the residuals of the ARIMA model, which would be the non-linear part of it.

※注意：1. 考生須在「彌封答案卷」上作答。

2. 本試題紙空白部分可當稿紙使用，試題須隨答案卷繳回。

3. 考生於作答時可否使用計算機、法典、字典或其他資料或工具，以簡章之規定為準。