

所別： 財務金融學系 碩士班 甲組(一般生)

科目： 統計

本科考試禁用計算器

*請在答案卷(卡)內作答

Part A: 問答與是非不定題

請依照題目順序以統計觀念說明理由或列舉推論過程回答下列問題。前五題為是非不定題，若僅回答對、錯、不一定而未予以說明解釋者一律不計分。若有題意不完整之處，請自行加註假設後作答。

1. 任何常態隨機變數的線性組合必定仍為常態分配。(5%)
2. 若護國神山公司董事長年終摸彩中，欲從 20 名優異表現的職工中抽出 5 位幸運的得主給予 100 張股票作為年終摸彩的特獎，已知籤筒中有五張紅色的籤與 15 張藍色的籤，若以取後不放回的方式抽獎，則第一位與第三位的抽籤獲獎率並不相同。(5%)
3. Given that $\{X_i\}_{i=1}^n$ are independently sampled from $N(2, 9)$, then Law of Large Number suggests that $\frac{\sum_{i=1}^n X_i^2}{n} \xrightarrow{n \rightarrow \infty} 4$. (5%)
4. Central Limit Theorem suggests that the sampling distribution of any estimator will eventually converge to Normality once the sample size is large enough. (5%)
5. The 95% confidence interval for population mean means that if you sample any sample observation from the population, 95 out of 100 times it will fall within this interval. (5%)
6. Suppose there were 100 bio-medical research teams globally that were mutually independently developing vaccine for COVID-19 virus around May 2020. If we apply t-test with significance level 5% individually to each of them under the null that the developed new vaccine is useless, what is the probability of falsely rejecting that at least one correct null collectively. (5%)
7. 神道銀行信用卡部發現他們透過社群軟體發送線上星巴克優惠券，約有兩成的社群軟體訂閱戶會下載使用。若該公司想擴大觸及率到至少有 1000 人使用該優惠券，請問他們至少應該要發出多少優惠券？(10%)
8. 試說明在一個存在多元變數 $\{X_1, X_2, X_3, X_4, X_5\}$ 的資料集中，你如何評估/計算給定 X_3, X_4, X_5 之下的 X_1 與 X_2 的條件相關係數，i.e., $\text{Corr}(X_1, X_2 | X_3, X_4, X_5)$? (10%)

Part B. 計算題及問答題

9. 張三欲反駁學歷無用論的說法，因此想研究學位種類是否對工作所得有顯著影響，打算收集 n 個社會新鮮人畢業後第一份工作薪水作為研究樣本，估計以下線性迴歸模型：

$$\text{Wage}_i = \alpha_0 + \alpha_1 \text{Degree}_i + \varepsilon_i,$$

其中 Wage_i 表第 i 個樣本每月薪資， $\text{Degree}_i=1$ 表第 i 個樣本的最高學位為高中以下， $\text{Degree}_i=2$ 表第 i 個樣本的最高學位為大學， $\text{Degree}_i=3$ 表第 i 個樣本的最高學位為碩士， $\text{Degree}_i=4$ 表第 i 個樣本的最高學位為博士，請評論這樣的研究方法，如認為不適當請提出你認為更適當的方法。(5%)

所別： 財務金融學系 碩士班 甲組(一般生)

科目： 統計

本科考試禁用計算器

*請在答案卷(卡)內作答

10. 估計此線性迴歸模型 $Y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \mu_i$ 後，欲檢驗 $H_0: \beta_1 = \beta_2 = 0$ ，請問是否只要 $H_0: \beta_1 = 0$ 及 $H_0: \beta_2 = 0$ 這兩種檢定結果有任一拒絕或同時拒絕即等於 $H_0: \beta_1 = \beta_2 = 0$ ，你是否同意此說法？請完整說明你的理由。(假設 β_1 的檢定量 t_1 與 β_2 的檢定量 t_2 的分配為獨立)(10%)

11. Michael wants to examine whether the gender of CEOs affects the probability of default in 2008. Michael collects 200 firm data and estimates the following the logit regression model:

$$\Pr(\text{Default}_{2008} = 1 | \text{Female}) = L(\alpha_0 + \alpha_1 \times \text{Female}_{i,2007}),$$

where $L(\cdot)$ is the logistic function; Default_{2008} is a dummy variable which equals one if firm i defaulted in 2008 and zero otherwise; $\text{Female}_{i,2007}$ is a dummy variable which equals one if the CEO of firm i is female in 2007 and zero otherwise. The results are shown below:

$$\Pr(\widehat{\text{Default}_{2008} = 1 | \text{Female}}) = L(0.6931 - 1.6740 \times \text{Female}_{i,2007})$$

Standard error (0.225) (0.311)

From the results, Michael argues that the presence of female CEOs would decrease the default probability of the firm by 1.674 percent. Please comment.

In addition, the odds ratio of *Female* is 0.1875. Please explain the meaning of this value and how do we get the value from the estimation results. (10%)

12. 請針對以下樣本進行線性迴歸分析，迴歸模型為：

$$Y_i = \alpha + \beta X + \varepsilon_i \text{ 且 } \varepsilon_i \overset{iid}{\sim} N(0, \sigma^2)$$

X	2	2	3	4	5	6
Y	4	6	6	7	7	9

請根據最小平方方法回答以下問題。(25%)

(一) 請估計 $\hat{\alpha}$ 與 $\hat{\beta}$ 以及 R^2 。

(二) 請根據迴歸分析計算過程完成以下 ANOVA 表格：

	Degree of freedom	Sum of Squares	Mean Square	F-stat
Regression	(a)	(d)	(g)	(i)
Residual	(b)	(e)	(h)	
Total	(c)	(f)		