

請將題號標示清楚，否則不允計分

壹、是非題(共三十分)：

正確標寫 T，錯誤標寫 F。(每題三分)

1. 當所有的資料值均為負數時，平均數、眾數、中位數及標準差亦均是負值。
2. Suppose that the probability of rain on a given day is 0.1 and the probability of my having a car accident is 0.005 on any day and 0.012 on a rainy day. The probability that on a given day it will rain and I will have a car accident is 0.0012.
3. Suppose the unemployment rate in the United States goes from 6.4% in one year to 5.6% in the next. The percentage change in the unemployment rate is -0.8.
4. 當平均數為零時，標準差亦等於零，但標準差等於零，平均數不一定為零。
5. 若分配是右偏分配，則平均數 < 中位數 < 眾數。
6. 平均數、中位數、眾數三者皆是用來測定一群資料分散度情況的統計測定數，而平均數總是優於中位數和眾數。
7. If X and Y are two independent random variables, then $\text{Var}(Y|X) = \text{Var}(Y)$ and $E(Y|X) = E(Y)$ and $\text{Cov}(X, Y) = 0$.
8. 假設變數 X 的母體分配未知，欲求母體平均數 μ 的 95% 信賴區間。現隨機抽取樣本數為 100 的一組樣本，得 $\bar{X} = 2$ ， $S_X^2 = 25$ ，則 μ 的 95% 信賴區間 = (1.02, 2.98)。我們可以說區間(1.02, 2.98)包含 μ 的機率為 0.95。
9. 信賴區間的寬度與準確度隨信賴水準的增加而增加。
10. 若 Z_X, Z_Y 分別是 X, Y 的標準化變數， X, Y 的相關係數等於 Z_X, Z_Y 的相關係數。

貳、解釋名詞(共三十分)

請簡單解釋下列統計名詞的意義(每題三分，只按照字面翻譯不給分)：

1. Contingency table
2. Coefficient of variation
3. Heteroscedasticity
4. Longitudinal data
5. Standard error
6. Spurious association
7. Binomial distribution
8. Regression sum of squares
9. p-value
10. The power of a hypothesis test

見背面

參、問答題(共佔四十分)：

一、以下為利用電腦進行迴歸分析(regression analysis)所呈現的結果。

regress educyr feducyr meducyr cage

Source	SS	df	MS			
Model	19749.8462	3	6583.28206	Number of obs = 1916		
Residual	17699.7629	1912	9.25719817	F(3, 1912) = 711.15		
				Prob > F = 0.0000		
				R-squared = 0.5274		
				Adj R-squared = 0.5266		
Total	37449.6091	1915	19.5559316	Root MSE = 3.0426		

educyr	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
feducyr	.3089197	.0221849	13.92	0.000	.2654105	.3524289
meducyr	.045937	.0247257	1.86	0.063	-.0025551	.0944292
cage	-.1184553	.0055994	-21.15	0.000	-.1294369	-.1074737
_cons	9.592141	.1492039	64.29	0.000	9.299521	9.88476

迴歸式可表為：

$$\widehat{educyr} = 9.59 + 0.309(feducyr) + 0.045(meducyr) - 0.1185(cage)$$

其中依變項 educyr 為受訪者的教育年數，

自變項為：

feducyr = 父親的教育年數

meducyr = 母親的教育年數

cage = (受訪者的年齡 - 全體受訪者的平均年齡)

請利用這個電腦分析的結果來回答以下問題：

- (1)請解釋 F(3, 1912) = 711.15 的意義。(五分)
- (2)請解釋 _cons 的係數 9.592141 的意義。(五分)
- (3)meducyr 的係數的 [95% Conf. Interval]的數值為 [-0.0025551, 0.0944292]，這是在檢定甚麼假設?檢定的結果如何?(五分)
- (4)你從這個迴歸模型中可以得到什麼結論?(五分)

二、以下交叉表為利用「台灣社會變遷基本調查」的全國代表性樣本所進行的分析。請從表格回答下列問題：

受訪者性別	是否為基督徒?		Total
	否	是	
男	1,882	80	1,962
女	1,921	90	2,011
Total	3,803	170	3,973

國立臺灣大學102學年度轉學生招生考試試題

題號： 33

科目：社會統計

題號： 33

共 3 頁之第 3 頁

- (1) 請問女生為基督徒的機率? (二分)
- (2) 請問男生基督徒佔所有樣本的比例(proportion)為? (三分)
- (3) 請問男生成為基督徒的勝算(odds)為? (二分)
- (4) 請問男生相對於女生成為基督徒的勝算比為(odds ratio)? (三分)
- (5) 請問母體中，男、女生基督徒的比例有沒有顯著的差異?請寫出計算過程與結論(五分)
- (6) 請用卡方檢定來檢定信仰基督教是否與性別有關?請寫出計算過程與結論(五分)



試題隨卷繳回