

類 科：衛生行政、食品衛生檢驗、衛生技術、養殖技術、海洋資源
科 目：生物統計學

考試時間：2小時

座號：_____

※注意：(一)可以使用電子計算器。

(二)不必抄題，作答時請將試題題號及答案依照順序寫在試卷上，於本試題上作答者，不予計分。

一、(一) 100 人的 ABO 及 Rh 血型如下表，請估計：在 Rh 陽性的條件下，A 型的機率。(5 分)

血型	A 型	B 型	AB 型	O 型
Rh 陽性	45	15	10	10
Rh 陰性	8	5	5	2

(二)心臟手術成功機率為 0.5，求 3 次手術中 3 次都成功的機率為何？(5 分)

(三)如果臺灣地區 20 歲女子體重為常態分布，其機率密度函數 (probability density function) 如下，請問體重的標準差是多少？(5 分)

$$f(x) = \frac{1}{\sqrt{50 \times \pi}} \exp\left[-\frac{(x-45)^2}{50}\right] \quad -\infty \leq x \leq \infty$$

(四) X_1, X_2, \dots, X_n 是從算數平均數為 μ 標準差為 σ 的常態分布群體中隨機抽出的樣本，則統計量 $\sum[(X_i - \bar{X})^2 \div \sigma^2]$ 的機率分布為何種分布？(5 分)

請以下列資料分析回答第二題至第四題：

在某地區進行有關心血管疾病的調查研究，研究者從該地區 40 歲以上的人群隨機選取 100 人，分別測量或記錄其身高(m)、體重(kg)、質量指數 ($BMI = kg/m^2$)、性別、年齡、有否吸菸、有否心血管疾病、日常活動量、血糖、膽固醇、收縮壓(mmHg)等變數。

機率分布之參考值：舉例說明： $t_{0.975, df=49} = 2.01$ 。t 表示 t 分布，下標 0.975 表示累積機率、df 表示自由度。2.01 則表示 t 分布橫座標的分位數值。

$$Z_{0.841} = 1.00 \quad Z_{0.90} = 1.28 \quad Z_{0.933} = 1.50 \quad Z_{0.95} = 1.645 \quad Z_{0.975} = 1.96 \quad Z_{0.977} = 2.00$$

$$\chi_{0.95, df=1}^2 = 3.84 \quad \chi_{0.975, df=1}^2 = 5.02 \quad t_{0.975, df=49} = 2.01 \quad t_{0.975, df=98} = 1.98 \quad F_{0.95, df1=1, df2=98} = 3.92$$

二、研究者想要比較男性與女性的質量指數平均值是否有顯著性的不同。質量指數資料經初步整理，結果如下所示：

性別	平均數	變異數	樣本數
男	25	6.0	50
女	23	4.0	50

(一)請估計男性群體質量指數平均值的百分之九十五信賴區間 (95%CI)。(5 分)

(二)寫出本研究雙尾檢定的虛無假設 (null hypothesis) 與對立假設 (alternative hypothesis)。(5 分)

(三)假定男、女兩群體質量指數的變異數相等，則群體質量指數變異數的估計值是多少？(5 分)

(四)假定男、女兩群體質量指數的變異數相等，統計檢定量的 t 值應是多少？(5 分)

(五)上述 t 檢定可視為變異數分析 (ANOVA) 的特例。如以變異數分析方式進行，請完成下列變異數分析表 (ANOVA table) 中(a)與(b)之數值。(每格 5 分，共 10 分)

變異來源	平方和(SS)	自由度(df)	均方(MS)	F 值
組間	XXXXXXXX	XXXXXXXX	XXXXXXXX	(b)
組內	XXXXXXXX	XXXXXXXX	(a)	XXXXXXXX
總和	XXXXXXXX	XXXXXXXX	XXXXXXXX	XXXXXXXX

(請接背面)

類 科：衛生行政、食品衛生檢驗、衛生技術、養殖技術、海洋資源
科 目：生物統計學

三、研究者想要了解年齡(X)與收縮壓(Y)的關係，希望能用年齡(自變數)來預測收縮壓(依變數)。年齡與收縮壓資料經初步整理，結果如下所示：

變數	樣本數	平均數	變異數	共變數(X, Y)
年齡(X)	100	55	25	25
收縮壓(Y)	100	130	400	

- (一)用年齡(X)來預測收縮壓(Y)的簡單直線迴歸模型(model)為 $\hat{y} = b_0 + b_1x$ 。請估計該模型的斜率(b_1)是多少？(5分)
- (二)請說明上述模型「斜率(b_1)」的含意。(5分)
- (三)如果將年齡(X)中心化，換言之 $X' = X - \text{mean}(X)$ 。用中心化年齡來預測收縮壓的簡單直線迴歸模型為 $\hat{y} = b_0 + b_1x'$ 。請說明該模型「截距(b_0)」的含意。(5分)
- (四)請估計該模型的決定係數(coefficient of determination)是多少？(5分)
- (五)迴歸分析中除須估計迴歸係數外也須呈現變異數分析表。請完成下列變異數分析表(ANOVA table)中(a)與(b)之數值。(每格5分，共10分)

變異來源	平方和(SS)	自由度(df)	均方(MS)	F 值
迴歸模型	XXXXXXXX	XXXXXXXX	XXXXXXXX	XXXXXXXX
殘差	XXXXXXXX	(a)	XXXXXXXX	XXXXXXXX
總和	XXXXXXXX	XXXXXXXX	(b)	XXXXXXXX

四、研究者想要了解性別、有否吸菸與有否心血管疾病間的關係。資料經初步整理，結果如下所示：

性別	吸菸	心血管疾病	人數
男性	是	是	20
男性	是	否	10
男性	否	是	10
男性	否	否	10
女性	是	是	2
女性	是	否	3
女性	否	是	5
女性	否	否	40

- (一)不考慮性別，請估計吸菸者對不吸菸者有心血管疾病的相對危險性(relative risk)。(5分)
- (二)不考慮性別，只討論有否吸菸與有否心血管疾病間的關係，進行卡方檢定時，如果不作連續校正(continuity correction)，則卡方值應是多少？(5分)
- (三)不考慮性別，只討論有否吸菸與有否心血管疾病間的關係，如改用不作連續校正的比例Z檢定(proportional Z test)，則統計檢定量Z值應是多少？(5分)
- (四)不考慮是否吸菸，請估計男性對女性有心血管疾病的勝算比(odds ratio)。(5分)