

113年第二次專門職業及技術人員高等考試營養師、護理師、社會工作師考試、
113年專門職業及技術人員高等考試心理師、法醫師、語言治療師、
聽力師、牙體技術師、公共衛生師考試、高等暨普通考試驗光人員考試試題

等 別：高等考試
類 科：公共衛生師
科 目：生物統計學
考試時間：2 小時

座號：_____

※注意：可以使用電子計算器。

甲、申論題部分：(50分)

- (一)不必抄題，作答時請將試題題號及答案依照順序寫在申論試卷上，於本試題上作答者，不予計分。
- (二)請以藍、黑色鋼筆或原子筆在申論試卷上作答。
- (三)本科目除專門名詞或數理公式外，應使用本國文字作答。

一、某研究團隊欲研究有無肥胖與慢性腎臟病的關係，收集 182 名慢性腎臟病病患及 178 名健康對照組，慢性腎臟病病患中 56 名有肥胖情形，健康對照組中 22 名有肥胖情形。請計算兩組間肥胖比率的差異的百分之九十五信賴區間，並請以統計方法檢定有無肥胖與慢性腎臟病間是否有統計顯著相關？設顯著水準 $\alpha=0.05$ 。(15 分)

(註： $Z_{0.995}=2.58$ ； $Z_{0.975}=1.96$ ； $Z_{0.95}=1.645$ ； $\chi^2_{1,0.99}=6.63$ ； $\chi^2_{1,0.975}=5.02$ ； $\chi^2_{1,0.95}=3.84$)

二、某研究團隊想要評估身體質量指數 (BMI) 與尿酸 (毫克/分升) 的關係，收集 249 名個案 (BMI 介於 15 至 35 之間)，得到簡單線性迴歸模式及變異數分析表如下：

$$\hat{y} = 5.124 + 0.069x$$

變異數分析表

變異來源	平方和	自由度	均方	F 值	P 值
迴歸模型	24.49	1	24.49	6.91	0.009
殘差	875.44	247	3.54		
總和	899.93	248			

請利用此簡單線性迴歸模式解釋自變項與依變項之間的關係？若欲檢定此迴歸模式是否具統計上的顯著性，其虛無假說及對立假說為何？請判斷此迴歸模式是否具統計上的顯著性？設顯著水準 $\alpha=0.05$ 。另請計算當身體質量指數為 25 時，其預期的尿酸值為何？(15 分)

三、某研究團隊想要評估有無規律運動與收縮壓間的相關性，他收集 125 名有規律運動者及 123 名無規律運動者，以比較有無規律運動的收縮壓平均值的差異，結果如下：

組別	人數	收縮壓平均值 (毫米汞柱)	收縮壓標準差 (毫米汞柱)
有規律運動組	125	122.8	19.2
無規律運動組	123	142.5	23.6

(一)請列出本問題之統計虛無假說及對立假說，並以適當的統計方法檢定有無規律運動組的收縮壓平均值是否有統計上的差異？設顯著水準 $\alpha=0.05$ 。(15分)

(二)當顯著水準改為 $\alpha=0.01$ 時，本題檢定結果的 p 值是否會改變？本研究的結論是否會改變？請解釋其原因。(5分)

(註： $t_{246, 0.95}=1.651$ 、 $t_{246, 0.975}=1.970$ 、 $t_{246, 0.990}=2.596$ 、 $t_{247, 0.95}=1.651$ 、 $t_{247, 0.975}=1.970$ 、 $t_{247, 0.990}=2.596$ 、 $t_{248, 0.95}=1.651$ 、 $t_{248, 0.975}=1.970$ 、 $t_{248, 0.990}=2.596$)

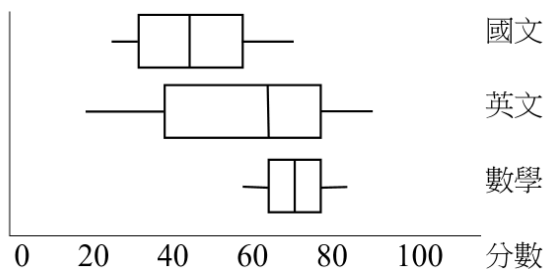
乙、測驗題部分：(50分)

代號：2108

(一)本試題為單一選擇題，請選出一個正確或最適當答案。

(二)共40題，每題1.25分，須用2B鉛筆在試卡上依題號清楚劃記，於本試題或申論試卷上作答者，不予計分。

1 下圖為某班期中考的國文、英文、數學分數的盒形圖，關於全距 (range)、中位數 (median)、四分位距 (interquartile range, IQR)、最大值、最小值，下列敘述何者正確？



- (A)數學的中位數最大，國文的 IQR 最小 (B)英文的全距最大，數學的 IQR 最小
(C)數學的最大值最大，英文的中位數最小 (D)國文的最小值最小，英文的 IQR 最大

2 某生物學家研究蘋果樹的枯萎病。他去一個大型蘋果園做調查，得到60%的蘋果樹開白花、其餘的開紫花；13%的蘋果樹感染枯萎病，而且10%開紫花的蘋果樹感染枯萎病。試問，開白花的蘋果樹感染枯萎病的機率？

- (A) 0.04 (B) 0.09 (C) 0.12 (D) 0.15

3 一母體平均數為 μ ，變異數為 σ^2 ，自母體抽取樣本 n ，有關中央極限定理 (Central Limit Theorem) 的敘述，何者錯誤？

- (A)不論隨機變數 X 在母群體為何種分布，只要其變異數不是無限大，當樣本數夠大時，樣本平均數 \bar{x} 的分布會趨近於常態分布
(B)樣本平均數分布的期望值 $\mu_{\bar{x}} = \mu$
(C)樣本平均數分布的標準差 $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$
(D)應用中央極限定理的先決條件是大樣本，所需樣本數大小與母群體分布的形狀無關

- 4 美國成年男性的體重分布近似於常態，平均為 170 磅、標準差 30 磅。隨機抽 25 位美國成年男性，其平均體重的 90% 上界為下列何者？
(註： $Z_{0.1}=1.282$, $Z_{0.05}=1.645$, $Z_{0.025}=1.960$, $Z_{0.01}=2.326$)
(A) 177.69 (B) 179.87 (C) 181.76 (D) 183.96
- 5 骨關節炎 (OA) 患者的肌肉功能以 10 分制來評估平衡協調性。隨機抽出 16 名 OA 患者，其平均分數 2.5、標準差 3.0。請計算 16 名 OA 患者平均分數的 95% 信賴區間之寬度？
(註： $Z_{0.05}=1.645$, $Z_{0.025}=1.960$, $t_{16, 0.05}=1.746$, $t_{16, 0.025}=2.120$, $t_{15, 0.05}=1.753$, $t_{15, 0.025}=2.131$)
(A) 2.630 (B) 2.940 (C) 3.180 (D) 3.197
- 6 關於檢定統計的檢定力 (power)，下列敘述何者正確？
(A) 相同條件下，資料的樣本數越大，檢定力越小
(B) 檢定力是不拒絕一個正確的虛無假說，所以越大越好
(C) 假設檢定的檢定力對應到診斷檢定的特異度 (specificity)
(D) 相同條件下，虛無假說與對立假說的抽樣分布重疊越少，檢定力越大
- 7 年齡介於 65~79 歲健康男性的血清尿酸的分布為近似於平均為 340 $\mu\text{mol/L}$ 、標準差為 80 $\mu\text{mol/L}$ 的常態分布。由此年齡層隨機抽出 4 位男性，其平均血清尿酸介於 260~420 $\mu\text{mol/L}$ 的機率約為下列何者？
(A) 0.69 (B) 0.73 (C) 0.87 (D) 0.95
- 8 某研究欲測試降血壓藥對男性、女性的效果是否相同，收集男性、女性各 30 位高血壓患者的服藥前收縮壓與服藥 5 週後收縮壓進行分析，下列何種方法最合適？
(A) 單一樣本 t 檢定 (one-sample t test)
(B) 配對樣本 t 檢定 (paired t test)
(C) 獨立樣本 t 檢定 (independent-sample t test)
(D) 皮爾森相關係數的 t 檢定 (t test for Pearson correlation coefficient)
- 9 高血壓研究的一個重要假設，是透過鈉攝取限制以降低血壓。其研究人員召集 8 個人參加一個鈉攝取限制研究，量測其原本的鈉攝取當作基線 (baseline) 及接受一週飲食諮詢之後的鈉攝取，資料如下表。以符號檢定 (sign test) 來檢定一週飲食諮詢是否能有效降低鈉攝取，檢定結果的 p 值為下列何者？
- | Person | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | Average | s.d. |
|------------|----|-----|----|----|----|----|----|----|---------|-------|
| Baseline | 14 | 12 | 22 | 14 | 18 | 42 | 16 | 14 | 19 | 9.798 |
| Week 1 | 10 | 24 | 16 | 20 | 16 | 32 | 12 | 14 | 18 | 7.709 |
| Difference | 4 | -12 | 6 | -6 | 2 | 10 | 4 | 0 | 1 | 7.010 |
- (A) 0.0078 (B) 0.1250 (C) 0.2266 (D) 0.4322
- 10 承上題，若以其他無母數方法來檢定一週飲食諮詢是否能有效降低鈉攝取，下列何者正確？
(A) 採 Wilcoxon 符號等級檢定 (Wilcoxon signed-rank test)，正等級和 $R_+ = 16.5$
(B) 採 Wilcoxon 符號等級檢定，正等級和 = 68
(C) 採 Wilcoxon 等級和檢定 (Wilcoxon rank-sum test)，其中一組等級和 $W_1 = 68$
(D) 採 Wilcoxon 等級和檢定，其中一組序等級 $W_1 = 16.5$

- 11 為研究一氧化碳暴露對於冠狀動脈疾病的影響，某研究從三個醫學中心招募男性參與者，首先進行基線特徵檢查，以確保來自不同中心的男性是可比較的。接著測量參與者的肺功能之第 1 秒用力呼氣體積 (FEV₁)，得到 A 中心 20 位參與者的平均 FEV₁ 為 2.63、標準差 0.49；B 中心 15 位參與者的平均 FEV₁ 為 3.10、標準差 0.52；C 中心 25 位參與者的平均 FEV₁ 為 2.83、標準差 0.50。若三個中心的男性群體平均 FEV₁ 分別為 μ_A 、 μ_B 、 μ_C ，以變異數分析 (ANOVA) 比較三個中心的男性平均 FEV₁，下列何者正確？
- (A) $H_0: \mu_A = \mu_B = \mu_C$ vs. $H_1: \mu_A \neq \mu_B \neq \mu_C$
 (B) 組間均方 (mean square between groups) 為 0.25
 (C) 組內均方 (mean square within groups) 為 0.7
 (D) F 檢定值為 2.8
- 12 某學生重複孟德爾的豌豆實驗，研究豌豆的兩種特徵：葉子形狀、豆莢顏色。葉子形狀由 A 基因座 (locus) 控制，AA 或 Aa 產生橢圓形葉子、aa 產生細長形葉子；豆莢顏色由 B 基因座控制，BB 或 Bb 產生綠色豆莢、bb 產生黃色豆莢。該學生以 AaBb × AaBb 的豌豆做實驗，得到子代包含 150 個橢圓形葉子且綠色豆莢、30 個橢圓形葉子且黃色豆莢、50 個細長形葉子且綠色豆莢、20 個細長形葉子且黃色豆莢。欲檢定 A、B 基因座是否獨立，則檢定值為下列何者？(註： $\chi^2_{1, 0.05} = 3.84$)
- (A) 2.536 (B) 3.972 (C) 4.464 (D) 5.718
- 13 關於變數 X 與 Y 的斯皮爾曼等級相關係數 (Spearman rank correlation coefficient, r_s)，下列何者錯誤？
- (A) 利用單調函數評量 X、Y 的相關性
 (B) r_s 的範圍在 -1 至 +1 之間
 (C) $r_s = -0.9$ 表 X、Y 相關性很低
 (D) $r_s < 0$ 表 X、Y 的相關方向相反，當 X 增加時，Y 趨向於減少
- 14 某研究從 520 名心血管疾病 (CVD) 男性和 1,100 名對照組男性抽取其血液樣本，並分析 200 個候選單核甘酸多態性 (single nucleotide polymorphisms, SNPs) 與 CVD 的關聯。每個 SNPs 編碼為 0、1、2。試問該採取什麼分析方法評估 SNPs 與 CVD 的關聯？
- (A) 皮爾森相關係數 (Pearson correlation coefficient)
 (B) 斯皮爾曼等級相關係數 (Spearman rank correlation coefficient)
 (C) McNemar 檢定
 (D) 卡方檢定 (chi-square test)
- 15 關於冠心病與年齡的研究。收集 14 位冠心病患者與 18 位沒有冠心病的人之年齡資料以羅吉斯迴歸 (logistic regression) 分析結果如下，下列敘述何者正確？

	B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I. for Exp(B)	
							Lower	Upper
Age (years)	0.130	0.047	7.749	1	0.005	1.139	1.039	1.248
Constant	-6.631	2.377	7.782	1	0.005	0.001		

- (A) 年齡每增加一歲，冠心病的風險 (risk) 增加 13.9%
 (B) 年齡每增加兩歲，冠心病的勝算比 (odds ratio, OR) 增加 29.7%
 (C) 係數的顯著性根據 Wald 檢定，Wald 是自由度為 1 的 t 分布
 (D) 因為 Exp(B) 的 95% 信賴區間不包含 0，所以相對應的 p 值小於 0.05

- 16 某個關於美國高中校園靜脈注射毒品的研究，收集了以下變項。毒品使用 (drug use)：1 代表在過去一年中有使用毒品，0 代表沒有使用毒品；是否有接受 HIV 教育 (HIV instruction)：1 (yes) 與 0 (no)；年紀 (age, 單位：歲)；性別 (gender)：1 (male) 與 0 (female)；具有 HIV 常識 (HIV knowledge)：1 (yes) 與 0 (no)，以羅吉斯迴歸分析使用毒品的機率並考慮交互作用項 (HIV instruction × gender) 得到以下結果。

Variable	Intercept	HIV instruction	age	gender	HIV knowledge	HIV instruction × gender
Coefficient	-1.18	0.04	-0.16	1.21	-0.19	-0.66
Std Error	0.86	0.42	0.09	0.42	0.05	0.51

(註：exp(1.21)=3.3535、exp(0.04)=1.0408、exp(-0.16)=0.8521、exp(-0.19)=0.8270、exp(-0.66)=0.5169、exp(-1.18)=0.3073、exp(-2.88)=0.0561)

一位 18 歲的女高中生無接受 HIV 教育、亦無 HIV 常識，試問她有使用毒品的機率？

- (A) 0.03274 (B) 0.02506 (C) 0.01695 (D) 0.00518
- 17 研究顯示，美國女性罹患乳癌的平均年齡比加拿大女性乳癌患者年輕。分析美國某醫院女性乳癌患者的年齡資料，檢定其平均年齡是否不同於 2020 年加拿大女性乳癌患者的平均年齡 (單位：歲) 如下表。假設該醫院的女性乳癌患者年齡之母體平均為 μ_A ， $\alpha=0.05$ ，下列敘述何者錯誤？

	Test value = 54					
	t	df	Sig.(two-tailed)	Mean difference	95% C.I. of the difference	
					Lower	Upper
Age (years)	6.224	1,206	< 0.001	2.388	1.64	3.14

- (A) 納入分析的女性乳癌患者共 1,206 人
 (B) 此檢定的對立假說為 $H_1: \mu_A \neq 54$
 (C) 該醫院的女性乳癌患者平均年齡為 56.388 歲
 (D) 檢定結果達統計上顯著
- 18 有關皮爾森 (Pearson) 相關係數和斯皮爾曼 (Spearman) 相關係數的敘述，下列何者最合適？
 (A) 都適合用來評估兩個次序變數 (ordinal variable) 之間的相關性
 (B) 當皮爾森相關係數為 0.25，統計檢定 p 值 < 0.0001，表示兩個變數之間有弱相關
 (C) 當資料有極端值 (extreme value) 時，使用皮爾森相關係數較為合適
 (D) 當皮爾森相關係數為 0，表示兩個連續變數之間不具有線性或其他形式的關聯
- 19 根據某職業棒球隊投手過去的經驗，練投 50 顆球，投進好球帶的次數為 40 顆。假設每次投球都是獨立的，倘若該投手今天練投 10 顆球，有 9 顆球以上投進好球帶的機率為何？
 (A) 0.375 (B) 0.107 (C) 0.268 (D) 4.198×10^{-6}
- 20 調查某醫院住院病人的住院日數，發現平均值是 10 天，中位數是 8 天，在全部住院病人中，住院 6 天的人數最多，資料為單峰分布。下列敘述何者正確？
 (A) 可能有一些病人住院特別久 (B) 標準差是 2 天
 (C) 住院日數符合常態分布 (D) 大約 50% 的病人住院少於 6 天
- 21 有關母群體參數信賴區間 (confidence interval, CI) 的寬度，下列敘述何者錯誤？
 (A) 樣本標準差越大，CI 寬度越大 (B) 樣本數愈大，CI 寬度愈小
 (C) 顯著水準 (α) 愈大，CI 寬度愈小 (D) 受樣本平均值的影響

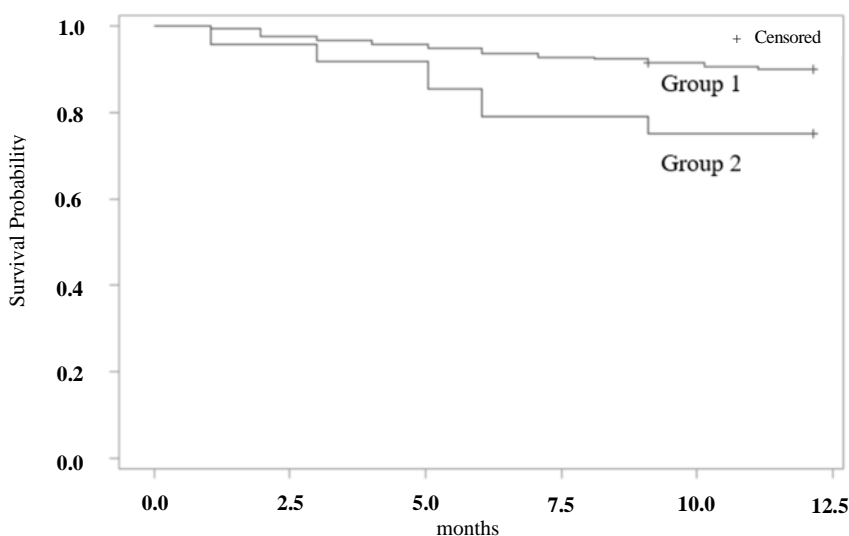
臨床試驗探討限時進食法（time-restricted eating）對體重（公斤）的影響，研究參與者為肥胖成人、年齡相似，分為兩組，A 組有額外採用限時進食策略，B 組則僅採用一般飲食控制。假定兩組的體重和體重變化皆為常態分布、母群體變異數相等。

	A. 限時進食+一般飲食 控制（49 人）	B. 一般飲食控制（49 人）
研究開始時的體重（還未採行介入策略）		
平均值	112.3	105.3
標準差	20.1	20.7
研究開始和結束時的體重變化		
平均值	-6.3	-4.0
95%信賴區間（雙尾）	-7.4, -5.2	-5.1, -2.9

（請回答第 22 至第 24 題）

- 22 若要檢定 A 組和 B 組的體重變化是否有統計上顯著差異，下列何者方法最適當？
- (A) Paired *t* test (B) Pearson correlation coefficient
(C) Two-sample *t* test (D) Two-sample test for binomial proportions
- 23 有關 A 組（限時進食+一般飲食控制）的體重變化，下列敘述何者正確？
- 說明： $t_{df, u}$ 表示 *t* 分布表的第 $u \times 100$ 個百分位；*df* 為自由度。 $t_{48, 0.975}=2.01$ 、 $t_{48, 0.95}=1.68$ 、 $t_{96, 0.975}=1.98$ 、 $t_{96, 0.95}=1.66$ 。
- (A) 體重變化標準差約為 0.55 公斤
(B) 體重變化標準差約為 3.83 公斤
(C) 體重變化標準差約為 4.58 公斤
(D) 根據現有的資訊，無法計算體重變化的標準差
- 24 有關 B 組（僅採用一般飲食控制）的體重變化，下列敘述何者正確？
- (A) B 組的體重變化，沒有達到統計上顯著意義
(B) 根據現有資訊，無法確定 B 組的體重變化是否達到統計上的顯著意義
(C) 若要進行統計檢定，了解一般飲食控制介入前後，B 組的體重是否有改變，虛無假設 (H_0) 為 $\mu_d=0$ ， μ_d 為試驗開始時和結束時的體重差的平均值
(D) 若要進行統計檢定，了解一般飲食控制介入後，B 組的體重是否有改變，虛無假設 (H_0) 為 $\mu_d = -4$ ， μ_d 為試驗開始時和結束時的體重差的平均值
- 25 以 *X* 代表自變數，*Y* 代表依變數，用線性迴歸來建立 *X* 和 *Y* 的關係，並利用 *X* 去預測 *Y*。下列何者不是其殘差項須滿足的前提假設？
- (A) 殘差呈現常態分布 (B) 殘差的平均數等於 0
(C) 殘差彼此相依 (D) 殘差的變異數具同質性 (constant variance)

- 26 關於假設檢定的 p 值，下列敘述何者正確？
- (A) p 值代表一個檢定的檢定力
 (B) p 值越小表示觀察到的資料越支持虛無假說
 (C) p 值是在虛無假說為真時，計算與觀察到的資料相等或更極端的機率
 (D) 相同條件下，單邊檢定的 p 值是雙邊檢定的 p 值之兩倍
- 27 臨床試驗探討接受 A、B 兩種治療藥物的病人，追蹤觀察病人存活的狀況。運用 Kaplan-Meier 法繪製存活曲線，如下圖，A 藥為 Group 1，B 藥為 Group 2，資料分析顯示在第 9 個月時，Group 1 存活率約為 91%，Group 2 為 75%。



下列敘述何者錯誤？

- (A) Kaplan-Meier 法屬於一種無母數統計方法 (non-parametric method)
 (B) 接受 A 藥治療的病人 (Group 1)，存活超過九個月的機率為 91%
 (C) 此組資料無法估計中位數存活時間 (median survival time)
 (D) 接受 B 藥治療組 (Group 2) 當中，最後一個有死亡個案的時間點約在第 12.5 個月
- 28 某都會區每日發生交通事故的平均次數為 3，已知每次交通事故的發生是獨立的，每日事故發生次數相當穩定，而且沒有季節和氣候的差異，事故發生次數也不受限。請計算交通事故次數的期望值和變異數分別為何？
- (A) 0.22、0.47 (B) 0.65、0.81 (C) 3、1.73 (D) 3、3
- 29 承上題，該都會區每日發生交通事故 4 次 (含) 以上的機率為何？

$$Pr(X = x) = \frac{e^{-\lambda} \lambda^x}{x!}, x=0, 1, 2, \dots, \infty, \lambda > 0$$

x	λ									
	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
0	.6065	.3679	.2231	.1353	.0821	.0498	.0302	.0183	.0111	.0067
1	.3033	.3679	.3347	.2707	.2052	.1494	.1057	.0733	.0500	.0337
2	.0758	.1839	.2510	.2707	.2565	.2240	.1850	.1465	.1125	.0842
3	.0126	.0613	.1255	.1804	.2138	.2240	.2158	.1954	.1687	.1404
4	.0016	.0153	.0471	.0902	.1336	.1680	.1888	.1954	.1898	.1755
5	.0002	.0031	.0141	.0361	.0668	.1008	.1322	.1563	.1708	.1755

- (A) 0.353 (B) 0.168 (C) 0.647 (D) 0.832

探討成人飲酒量（杯／週）和高密度脂蛋白膽固醇（HDL）（mmol/L）的關係，迴歸模式分析結果如下（顯著水準=0.05）

變數	自由度	參數估計值（係數）	標準誤	95%信賴區間
截距	1	1.2683	0.0101	1.2484, 1.2882
飲酒量	1	0.0059	0.0004	0.0051, 0.0067

R 平方=0.096

（請回答第 30 至第 32 題）

30 假設 y 表示 HDL 濃度， x 表示飲酒量。請問直線迴歸方程式為何？

- (A) $\hat{Y} = 1.2683 + 0.0101X$ (B) $\hat{Y} = 1.2683 + 0.0059X$
(C) $\hat{Y} = 0.0101 + 0.0004X$ (D) $\hat{Y} = 0.0101 + 1.2683X$

31 有關分析結果的解釋，下列敘述何者正確？

- (A) 整體的 HDL 平均值是 1.2683 mmol/L
(B) 皮爾森相關係數（Pearson correlation coefficient）是 0.096
(C) 皮爾森相關係數是 0.31
(D) 飲酒量和 HDL 的關聯未達到統計上顯著意義

32 若將 HDL 的單位轉換為 mg/dL， $1 \text{ mmol/L} = 18 \text{ mg/dL}$ ，再進行迴歸分析探討喝酒量和 HDL 的關聯，設原本（未轉換以前）迴歸係數為 β 、喝酒量和 HDL 的皮爾森相關係數為 r ，經單位轉換後數值會如何改變？

- (A) 迴歸係數和相關係數都不變 (B) 迴歸係數變為 $\beta \times 18$ 、相關係數不變
(C) 迴歸係數不變、相關係數變為 $\beta \times 18$ (D) 迴歸係數和相關係數都變為 $\beta \times 18$

33 收集 90 個國家的資料，探討兒童的含糖飲食攝取量（公斤／人／年）和齲齒指數的關係，齲齒指數高表示齲齒顆數越多。設 y 為齲齒指數、 X_1 為含糖飲食攝取量、 X_2 為國家開發程度（低設為 1、高設為 0）。得到線性迴歸模式為 $\hat{Y} = 1.68 + 0.04X_1 - 0.35X_2$ ，決定係數（coefficient of determination）為 0.23。假設迴歸係數都達到統計顯著意義。下列敘述何者正確？

- (A) 在齲齒指數的總變異量中，能被含糖飲食攝取量所解釋的變異量百分比約為 23%
(B) 控制含糖飲食攝取量後，與高度開發國家的兒童相比，低度開發國家兒童的平均齲齒顆數較多
(C) 國家開發程度的迴歸係數比含糖飲食攝取量的迴歸係數大，表示國家開發程度的重要性比較大
(D) 含糖攝取量 4 公斤／人／年的高度開發國家，兒童平均齲齒指數估計約為 1.84 顆

- 34 一個 1:1 病例對照研究探討暴露到有機氯農藥和乳癌的關聯，假如使用雙尾檢定，顯著水準為 5%，統計檢定力 (power) 為 90%，要偵測到勝算比 (odds ratio) 為 2 的差異，病例組和對照組各約需 230 人。下列何者改變，會增加研究所需要的樣本數？
- (A) 統計檢定力設為 80% (B) 顯著水準改為 1%
(C) 改為單尾檢定 (D) 欲偵測的勝算比改為 2.5
- 35 一研究評估低濃度的一氧化碳 (CO) 暴露對心臟病患者運動耐受性的影響，共 15 位病人加入研究，每人皆分別暴露於一氧化碳和一般空氣，進行同樣的運動，直到出現心絞痛、疲勞或高血壓等狀況，記錄運動持續時間，如下：

病人編號	運動持續時間 (秒)	
	暴露於低濃度一氧化碳	暴露於一般空氣
1	285	270
2	220	220
3	370	406
4	480	520
5	1020	1020
6	540	540
7	180	180
8	780	790
9	840	860
10	330	505
11	360	540
12	1010	1020
13	755	710
14	1170	1280
15	750	615

若要進行統計檢定，下列何種方法最合適？

- (A) Kruskal-Wallis test (B) Wilcoxon rank sum test
(C) Wilcoxon signed rank test (D) McNemar's test
- 36 想了解心血管疾病家族史的青少年之平均膽固醇濃度是否高於一般青少年，若收集資料後進行統計檢定，顯著水準 α 為 0.05，結果發現統計檢定力 (statistical power) 為 24%，下列何者正確？
- (A) 在虛無假設成立之下，拒絕虛無假設的機率是 24%
(B) 在對立假設成立之下，拒絕虛無假設的機率是 24%
(C) 把顯著水準 α 降低，可提高統計檢定力
(D) 改變取樣設計，讓樣本的標準差變大，可以增加統計檢定力

- 37 探討和心臟病有關的健康行為因子，運用羅吉斯迴歸 (logistic regression) 分析，年齡為連續值 (不分組)，抽菸史分為三組 (從不抽菸、已戒菸、目前有抽菸)，身體活動量分為兩組 (低度、中或強度)，得到以下結果 (設統計顯著水準為 0.05)：

變項	勝算比 (odds ratio)	95% 信賴區間
年齡 (歲)	1.044	1.015, 1.074
抽菸史		
已戒菸 (與從不抽菸相比)	1.345	0.772, 2.343
目前有抽菸 (與從不抽菸相比)	2.160	1.339, 3.484
身體活動量		
中或強度 (和低度相比)	0.947	0.655, 1.370

有關年齡的羅吉斯迴歸係數，下列敘述何者正確？

- (A) 是 $\ln(1.044)$ (B) 是 1.044
(C) 是 $e^{1.044}$ (D) 統計檢定的虛無假設是 1
- 38 承上題，下列敘述何者正確？
- (A) 控制年齡、抽菸史以後，相較於身體活動量低者，活動量中或強度者比較不會患心臟病，有達到統計上顯著意義
(B) 羅吉斯迴歸係數之統計檢定採用 t test
(C) 若檢定抽菸史這個變數整體是否有統計上顯著意義，需要設定 2 個虛擬變數 (通常設為 1 或 0)，統計檢定的自由度是 3
(D) 根據羅吉斯迴歸公式，可得到每個研究對象患心臟病的估計機率
- 39 探討女性生產年齡是否和罹患乳癌有關，從多家醫院收集資料，將第一胎生育年齡分為 <30 歲和 ≥30 歲，發現在 1000 位乳癌病人、5000 位沒有乳癌的病人當中，第一胎生育年齡 ≥30 歲者分別有 210 人、700 人。若進行統計分析，下列何種檢定方法最不適合？
- (A) Chi-square test (B) Two-sample test for binomial proportion
(C) Spearman correlation (D) Logistic regression
- 40 運用羅吉斯迴歸 (logistic regression) 探討疾病 Y 的危險因子， Y 為兩分類的依變數 (有或無患病)，危險因子為自變數 X 。下列敘述何者正確？
- (A) 假如只有一個類別型自變數 X ，使用羅吉斯迴歸和列聯表分析 (contingency-table analysis) 檢定 Y 跟 X 的關聯性，會得到相似的結果
(B) 可估計出勝算比，其意義為：患有 Y 的機率對沒有罹患 Y 的機率的比值
(C) 前提假設之一是連續型的自變數之間呈線性關聯
(D) 羅吉斯迴歸模型的估計是採用最小平方法 (ordinary least squares)