

111年第二次專門職業及技術人員高等考試中醫師考試分階段考試、營養師、護理師、社會工作師考試、111年專門職業及技術人員高等考試法醫師、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、某研究者欲探討某空氣污染物的濃度是否與某呼吸道疾病之急診就診人數有關，他在某醫院觀察了 10 天，得到資料如下：

天數	1	2	3	4	5	6	7	8	9	10
空氣污染物之濃度	72	112	118	124	137	166	148	137	87	99
呼吸道疾病之急診人數	41	97	95	123	85	128	132	114	37	48

試計算：

- (一)空氣污染物 (X) 對於呼吸道疾病急診人數 (Y) 之迴歸方程式。(10 分)
- (二)此迴歸方程式之決定係數，並說明其意義為何。(10 分)

二、某研究者欲評估某藥劑在不同年齡層之症狀緩解時間長短，他在三個不同的年齡層各隨機抽取了 11 位男性 (總樣本數：33 位)，受訪者自述之症狀緩解時間如下：

青年人：23 ± 12 分

壯年人：27 ± 10 分

中年人：31 ± 8 分

試檢定三個年齡層之症狀緩解時間是否有差異 ($\alpha=0.05$)。(15分)

$F_{0.95,2,30}=3.32$ 、 $F_{0.975,2,30}=4.18$ 、 $F_{0.95,2,31}=3.30$ 、 $F_{0.975,2,31}=4.16$ 、

$F_{0.95,3,30}=2.92$ 、 $F_{0.975,3,30}=3.59$ 、 $F_{0.95,3,31}=2.91$ 、 $F_{0.975,3,31}=3.57$

三、某研究者欲比較三種不同廠牌疫苗對兒童產生副作用之情況是否有差異，他從某醫院隨機抽取了 200 位兒童，得到結果如下，試以 $\alpha=0.05$ 來進行檢定。(15 分)

	產生副作用的人數	未產生副作用的人數
廠牌A	28	39
廠牌B	38	30
廠牌C	34	31

$$\chi^2_{0.90,1}=2.71、\chi^2_{0.95,1}=3.84、\chi^2_{0.975,1}=5.02、\chi^2_{0.90,2}=4.61、\chi^2_{0.95,2}=5.99、\chi^2_{0.975,2}=7.38$$

$$\chi^2_{0.90,3}=6.25、\chi^2_{0.95,3}=7.81、\chi^2_{0.975,3}=9.35、\chi^2_{0.90,6}=10.64、\chi^2_{0.95,6}=12.59、\chi^2_{0.975,6}=14.45$$

乙、測驗題部分：(50分)

代號：2110

(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 某醫學中心為了研究以 β 類澱粉蛋白質 (amyloid β 42) 作為阿茲海默症 (Alzheimer's disease) 的生物指標是否適合，特針對某縣市居民進行調查 (已知阿茲海默症盛行率為 0.02)。該醫學中心由此縣市中找了 300 位阿茲海默症患者及 700 位健康者進行測試，所得結果如下表。從此縣市隨機抽取一民眾，若 β 類澱粉蛋白質呈陽性反應，試問其得阿茲海默症的機率 $Pr(D^+|T^+)$ 最接近數值為何？

檢驗結果	阿茲海默症患者 (D^+)	健康者 (D^-)	合計
陽性 (T^+)	280	112	392
陰性 (T^-)	20	588	608
合計	300	700	1000

- (A) 0.872 (B) 0.998 (C) 0.106 (D) 0.632
- 2 自由度為 1 的卡方分布機率密度函數，右尾 5% 的 X 軸臨界值 (閾值)，與下列何者臨界值相同？
 (A)標準常態分配的右尾 2.5% (B)標準常態分配的右尾 5%
 (C)標準常態分配的右尾 2.5%之平方值 (D)標準常態分配的右尾 5%之平方值
- 3 石虎為瀕臨絕種之野生動物，主要棲息環境為淺山地區。由於近年來淺山地區受到人類大量開發，石虎在道路上遭到車輛輾斃之事件【稱為路殺 (road kill)】時有所聞。已知每次路殺的發生是獨立的、路殺發生次數也不受限。若臺灣某路段每年發生路殺的平均次數為 48，若令 Y 表示該路段每月發生路殺的次數，則 Y 分布之變異數為何？
 (A) 48 (B) 4 (C) 12 (D) 6

8 單因子變異數分析表如下，請問 X 之值為何？

變異量來源	自由度	平方和	均方	F 值
組間	2	17692	X	16
組內	51	28091	551	
總和	53	45783		

- (A) 8846 (B) 17692 (C) 5897 (D) 347

9 研究者欲了解服用避孕藥丸後對 SBP 的值是否會有影響，今有 15 位婦女進入此研究，此 15 位婦女從未服用過避孕藥，研究者在給服藥之前，先測得每位婦女的 SBP 值，給服藥兩週後再測得婦女的 SBP 值，以茲比較，所得的資料如下表。試檢定服用避孕藥對 SBP 值是否有影響，下列何種方法適用？

婦女代號 (i)	服藥前 SBP	服藥後 SBP
1	115	128
2	112	115
3	107	106
4	119	128
5	115	122
6	138	145
7	126	132
8	105	109
9	104	102
10	115	117
11	105	118
12	102	115
13	111	106
14	113	124
15	115	119

- (A) 兩組獨立樣本 t 檢定 (B) 成對樣本 t 檢定
(C) 兩組獨立樣本 Z 檢定 (D) 單一樣本 Z 檢定

10 若使用變異數分析並以 0.05 的型一誤差 (type I error rate) 比較某一所高中的 5 個班級 (A, B, C, D, E) 的國文成績，檢定結果是達統計上顯著性差異，因此如果想要進一步進行所有兩兩班級國文成績的多重比較，使用 Bonferroni 法則，此時型一誤差應該要調整成下列那個數值？

- (A) 0.005 (B) 0.01 (C) 0.025 (D) 0.0025

- 11 分析 40 位 12~17 歲青少年的年齡 (Age, 歲) 與身高 (公分) 的關聯性分析, 得到下表, 下列敘述何者錯誤?

Variable	Estimate	Lower bound of 95% CI	Upper bound of 95% CI
Intercept	97.03	71.74	122.33
Age (歲)	4.18	2.37	6.00
Male (男生=1, 女生=0)	6.17	0.86	11.49

$R^2 = 0.45$

- (A) 決定係數是 0.45
 (B) 校正性別後, 年齡每增加 1 歲身高增加 4.18 公分
 (C) 校正年齡後, 男生比女生高 6.17 公分
 (D) 校正年齡後, 女生比男生矮, 但未達顯著性差異
- 12 下列資料為二株新冠病毒在蝙蝠體內存活天數, 此資料屬小樣本且兩個母群體不符合常態分布, 若要以此資料檢定二株新冠病毒在蝙蝠體內存活天數是否有統計上顯著差異, 請問應該使用何種檢定方法?

存活天數 (days)	
H 型	N 型
53	46
56	52
60	58
60	59
78	77
87	117
102	120
117	123
120	134
121	156
	160

- (A) Signed test
 (B) Wilcoxon signed-rank test
 (C) Wilcoxon rank sum test
 (D) Kruskal-Wallis test
- 13 一個病例對照研究結果如下表。原始研究的樣本數為 70 人, 假設樣本數變成 10 倍, 共 700 人, 請問下列何者最可能變大?

	Disease	No Disease
Exposure	10	5
No Exposure	20	35

- (A) 勝算比 (Odds Ratio)
 (B) 檢定統計量 (Test statistic)
 (C) P 值 (P-value)
 (D) 95% 信賴區間 (95% confidence interval)

- 14 若比較肺癌病人的診斷細胞型態之 Karnofsky Performance Status (KPS) 得到單因子變異數分析表結果如下，以 0.05 的型一誤差 (type I error rate) 來判定，請問下列敘述何者正確？

變異量來源	自由度	平方和	均方	F 值	P 值
組間	3	3740.483	1246.83	2.9704	0.0341
組內	137	57505.474	419.75		
總和	140	61245.957			

- (A)此分析資料應該有 4 種細胞型態及 141 筆資料，這幾種細胞型態的 KPS 平均值不完全相等
 (B)此分析資料應該有 4 種細胞型態及 141 筆資料，這幾種細胞型態的 KPS 變異數不完全相等
 (C)此分析資料應該有 4 種細胞型態及 140 筆資料，這幾種細胞型態的 KPS 平均值不完全相等
 (D)此分析資料應該有 4 種細胞型態及 140 筆資料，這幾種細胞型態的 KPS 變異數不完全相等
- 15 某大醫院的醫師正從事於頭痛與高血壓的關聯性研究，他發現 273 位患有高血壓的勞工中，共有 26 位常常有頭痛的情形，而 1046 位未罹患高血壓的勞工中，有 44 位常常有頭痛的情形，今欲以此資料推論頭痛與高血壓是否有關聯，則下列檢定方法何者正確？
 (A)卡方檢定 (B) ANOVA 檢定
 (C)獨立樣本 t 檢定 (D)配對 t 檢定
- 16 分析 2×2 列聯表資料時，下列何種情況不適合使用卡方檢定？
 (A)某個細格的觀察次數小於 5 (B)某個細格的期望次數小於 5
 (C)某個細格的觀察次數小於 10 (D)某個細格的期望次數小於 10
- 17 假設隨機變數 X 代表病人服用抗病毒藥品的天數，此母群體呈右偏分布，其平均值為 μ 且標準差為 σ ，如果依據中央極限定理，則下列何者在樣本數增加時會接近常態分布？
 (A)隨機樣本 X_i (B)樣本平均值 \bar{X}
 (C)樣本變異數 $\sum (X_i - \bar{X})^2 / (n-1)$ (D)標準化隨機樣本 $(X_i - \mu) / \sigma$
- 18 針對某個樣本數為 20 的樣本，所計算的標準差 (standard deviation) 與標準誤 (standard error)，下列敘述何者正確？
 (A)標準差 = 標準誤 $\times \sqrt{20}$ (B)標準誤 = 標準差 $\times \sqrt{20}$
 (C)標準差 = 標準誤 $\times 20$ (D)標準誤 = 標準差 $\times 20$
- 19 若針對某一社區居民隨機抽樣 300 人，此樣本之平均年齡為 55 歲，已知其平均年齡顯著不同於 64 歲，假設社區居民的年齡呈常態分布，則下列何者比較可能是平均年齡的雙尾 95% 信賴區間？
 (A) (57, 71) (B) (48, 62) (C) (54, 74) (D) (45, 65)
- 20 根據 COVID-19 全球疫情地圖，得知臺灣疫苗接種覆蓋率 86.79%，如果隨機在臺灣抽取一萬人，計算此樣本的疫苗接種覆蓋率，此樣本觀察到接種疫苗人數的變異數為下列何者？
 (A) 1146 (B) 1321 (C) 7532 (D) 8679
- 21 針對統計檢定力 (statistical power, $1-\beta$) 的敘述，下列何者較正確？
 (A)當虛無假設為真的，卻推翻虛無假設的機率
 (B)當虛無假設是真的，卻沒有推翻虛無假設的機率
 (C)當對立假設是真的，卻推翻虛無假設的機率
 (D)當對立假設是真的，卻沒有推翻虛無假設的機率

- 22 某個研究罹患癌症的危險因子分析結果如下表所示，以 0.05 的型一誤差 (type I error rate) 來判定，下列敘述何者錯誤？

		勝算比 odds ratio	95% 信賴區間
Gender	male	0.94	(0.64, 1.37)
	female	1.00	
Age (years)	age < 40	1.00	
	40 ≤ age < 60	1.52	(0.96, 2.41)
	60 ≤ age	2.21	(1.28, 3.80)

- (A) 男性罹患癌症的危險性較女性高
 (B) 40 至 60 歲的人相較於 40 歲以下罹患癌症的危險性未達統計上顯著性
 (C) 年紀 60 歲以上者相較於 40 歲以下罹患癌症的危險性較高
 (D) 這個表格的結果是使用羅吉斯迴歸 (logistic regression) 計算而得
- 23 一項探討醫學檢驗值與胰臟癌關係的病例對照研究，研究過程中收集了飯前血糖、飯後血糖、糖化血色素及總膽固醇，若想要了解有或沒有胰臟癌之相關醫學檢驗值互相校正後，何項相關醫學檢驗值較有顯著意義，下列何種統計方法最合適？
 (A) 簡單線性迴歸 (B) 複線性迴歸 (C) 簡單羅吉斯迴歸 (D) 複羅吉斯迴歸
- 24 一組 100 人的隨機樣本，進行暴露 (X) 與疾病 (Y) 的關係研究，變數設定以及研究結果整理為下列 2 × 2 的列聯表，如果以羅吉斯迴歸分析之後，截距 β_0 的估計值為何？

	有病 (Y = 1)	沒病 (Y = 0)
有暴露 (X = 1)	10	20
無暴露 (X = 0)	40	30

- (A) 10/40 (B) 20/30 (C) $\log(10/20)$ (D) $\log(40/30)$
- 25 利用羅吉斯迴歸分析有無用藥 (Treatment, yes vs. no) 對於疾病的治癒效果 (Cure, yes vs. no)，發現在男性族群中，有用藥物治療的病患勝算 (odds)，是安慰劑病患勝算的 2 倍，但是女性族群中，勝算比降為 1.2 倍。請問這個研究中，對於性別變數的敘述何者最適切？
 (A) 性別變數是干擾因子 (confounder)
 (B) 性別變數修飾了藥物與治癒的關係 (effect modifier)
 (C) 性別變數只是共變數 (covariate)
 (D) 性別變數可以預測疾病
- 26 想研究高中生男女的體脂肪有無差異，隨機收集了年齡相近的 13 位男生與 11 位女生，描述統計如下所示。欲以 two sample t-test 檢定，若假設兩組變異數相等，此變異數 (pooled variance) 的估計值最接近之數值為何？

組別	樣本數 (n)	平均數 (\bar{x})	標準差 (s)
女同學	11	22	5
男同學	13	15	7

- (A) 35 (B) 36 (C) 37 (D) 38

27 分析 40 位 12~17 歲青少年的年齡 (Age) 與身高 (公分) 的關聯性分析，得到下表，下列敘述何者正確？

Variable	Estimate	Lower bound of 95% CI	Upper bound of 95% CI
Intercept	96.82	70.09	123.56
Age	4.44	2.54	6.34

$$R^2 = 0.37$$

- (A)皮爾森相關係數 (Pearson correlation coefficient) 是 0.37，迴歸係數 (regression coefficient) 是 4.44
 (B)皮爾森相關係數是 0.61，迴歸係數是 4.44
 (C)皮爾森相關係數是 0.61，迴歸係數是 96.82
 (D)身高有 37% 可以被年齡所解釋，迴歸係數是 96.82
- 28 分析年齡 (age)、性別 (gender)、教育程度 (educ) 與美國總統選舉 (vote) 的相關性，這些變數的數值編碼 (coding) 與分析結果如下。請問男性對比女性選民 (Male vs. Female)，投給 Trump (vote = 1 vs. 0) 的勝算比為何？

vote

0 = ' Clinton '

1 = ' Trump ' ;

gender

1 = ' Male '

2 = ' Female ' ;

educ

1 = ' HS Not Completed '

2 = ' Completed HS '

3 = ' College <4 Years '

4 = ' College 4 Year Degree '

5 = ' Advanced Degree ' ;

Analysis of Maximum Likelihood Estimates						
Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	0.2754	0.1959	1.9759	0.1598
gender	Female	1	-0.3559	0.0841	17.8915	<.0001
gender	Male	0	0	-	-	-
age		1	0.0131	0.00246	28.3128	<.0001
educ		1	-0.2518	0.0387	42.3355	<.0001

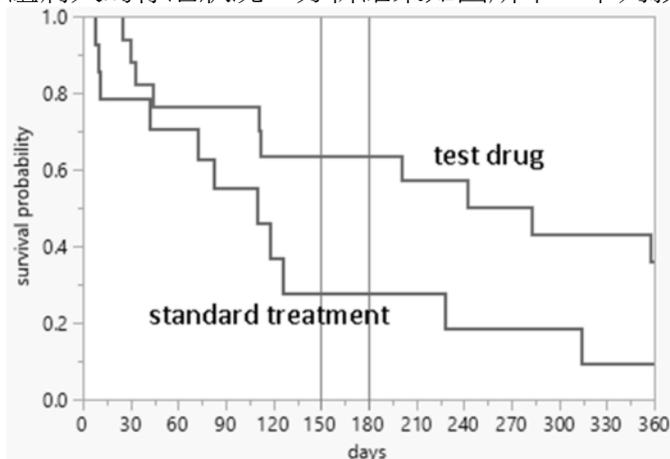
(A) 0.3559

(B) -0.3559

(C) $e^{-0.3559}$

(D) $e^{0.3559}$

- 29 某個治療肺癌病人的臨床試驗比較試驗用藥 (test drug) 和標準治療 (standard treatment) 兩組病人的存活狀況，分析結果如圖所示，下列敘述何者錯誤？



- (A) 此存活曲線是由 Kaplan-Meier 估計值 (estimates) 所繪製而成
 (B) 試驗用藥 (test drug) 比標準治療 (standard treatment) 的存活狀況好
 (C) 試驗用藥 (test drug) 的中位數存活時間在 180 天以上
 (D) 標準治療 (standard treatment) 的中位數存活時間在 150 天以上
- 30 在進行比較實驗組與對照組之衛教介入前後知識分數的改變量平均值時，統計分析軟體提供了下列的結果，以 0.05 的型一誤差 (type I error rate) 來判定：
 檢定兩組變異數是否有差異的 F 檢定：p-value = 0.010
 Equal variance 之 t 檢定：p-value = 0.001
 Unequal variance 之 t 檢定：p-value = 0.212
 請問下列敘述何者正確？
 (A) 實驗組和對照組的改變量平均值達顯著性差異，此檢定的 p 值是 0.010
 (B) 實驗組和對照組的改變量平均值達顯著性差異，此檢定的 p 值是 0.001
 (C) 實驗組和對照組的改變量平均值沒有達顯著性差異，此檢定的 p 值是 0.212
 (D) 實驗組和對照組的改變量平均值沒有達顯著性差異，此檢定的 p 值是 0.001
- 31 在執行存活分析時需要定義每筆個案的資料之追蹤結果是屬於事件 (event) 資料，還是設限 (censor) 資料，某一結腸直腸癌的標靶藥品治療研究，病人在試驗結束後所登錄的追蹤結果分別是：A=存活，B=死於肺癌，C=死於結腸直腸癌，D=死於肺癌及結腸直腸癌以外的其他癌症，E=死於非癌症的其他原因，因此在計算各種死亡率時，下列敘述何者錯誤？
 (A) 計算癌症存活率時，BCD 設定成事件資料，AE 設定成設限資料
 (B) 計算結腸直腸癌存活率時，C 設定成事件資料，ABDE 設定成設限資料
 (C) 計算總存活率時，BCDE 設定成事件資料，A 設定成設限資料
 (D) 計算肺癌存活率時，BC 設定成事件資料，ADE 設定成設限資料
- 32 一個研究要分析 6 所大學的資料，想探討一年級新生的 BMI 數值是否相等，於是在每一間大學隨機抽取 100 人進行試驗，最終選定變異數分析 ANOVA 作為分析工具。此方法的檢定統計量為下列何者 (F 是累積機率密度函數， α 是顯著水準)？
 (A) $F_{0.05}(5, 594)$ (B) $F_{0.05}(5, 599)$ (C) $F_{0.95}(5, 594)$ (D) $F_{0.95}(5, 599)$
- 33 一個預測變數有 3 個類別，暴露等級分別為輕度、中度、重度。此變數在線性迴歸分析中，欲以虛擬變數的方式分析，請問要生成幾個虛擬變數？
 (A) 1 (B) 2 (C) 3 (D) 4

34 若 A 和 B 兩事件發生機率已知分別為 $P(A) = 0.4$ ； $P(B) = 0.6$ ； $P(A|B) = 0.4$ ，試求 $P(A \cup B)$ 為何？

- (A) 0.76 (B) 1.00 (C) 0.24 (D) 0.20

35 一個研究探討新發現的兩個發炎指數之間的相關性，已知其中一個發炎指數為指數分配，另一個為 F 分配 $F(5, 2)$ ，下列何種統計方法最適用於此研究？

- (A) Spearman correlation coefficient (B) Pearson correlation coefficient
(C) Signed rank test (D) Mann-Whitney U test

36 Cox proportional hazards model 如下，其中 Age 的單位為歲 (years)

$$h_i(t) = h_0(t) \times e^{0.1 \times \text{Age} + 0.2 \times \text{BMI} + 0.3 \times \text{Male}}$$

對於相同 BMI 值的男性，請問 30 歲的人對比 20 歲 (Age = 30 vs. Age = 20) 的風險 (hazard) 的比值是多少？

- (A) 0.1 (B) e^1 (C) $e^{0.1}$ (D) $10 \times e^{0.1}$

37 假設 A 為血壓器測量顯示有高血壓症狀， \bar{A} 則為血壓器測量顯示正常值 (無高血壓症狀)，B 為測試者有高血壓， \bar{B} 則為正常測試者 (血壓值正常)，請問此血壓器的特異度 (specificity) 是下列那一個條件機率？

- (A) $P(A|B)$ (B) $P(\bar{A}|\bar{B})$ (C) $P(B|A)$ (D) $P(\bar{B}|\bar{A})$

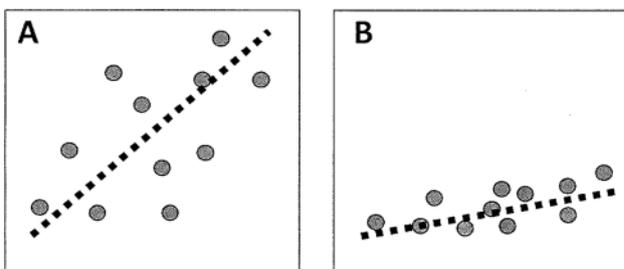
38 某項研究調查罹患口腔癌的危險因子，經由羅吉斯迴歸計算出抽菸與嚼食檳榔的勝算比和 95% 信賴區間，分別為抽菸 1.5 (0.9~2.5) 及嚼食檳榔 2.5 (1.5~4.1)，若兩個危險因子各自以 0.05 的型一誤差 (type I error rate) 來判定，則下列敘述何者正確？

- (A) 抽菸與嚼食檳榔都是達統計上顯著性的危險因子
(B) 僅嚼食檳榔是達統計上顯著性的危險因子
(C) 僅抽菸是達統計上顯著性的危險因子
(D) 抽菸與嚼食檳榔都沒有達統計上顯著性

39 某位老師研究 2000 位新生兒，結果發現新生兒出生體重 (y，以公克為單位) 可能與媽媽生產年齡 (x，連續變數，以歲為單位) 有關，該老師決定採用 simple linear regression 以 x 來預測發生 y 的可能性，可得到迴歸方程為 $y = \hat{\alpha} + \hat{\beta}x$ 。已知 $\hat{\beta} < 0$ 且其 $p < 0.05$ ，請問下列敘述何者正確？

- (A) 當媽媽生產年齡愈大，其新生兒出生體重愈輕
(B) 當媽媽生產年齡愈大，其新生兒出生體重愈重
(C) 無法判定媽媽生產年齡與其新生兒出生體重的關係
(D) 媽媽生產年齡與其新生兒出生體重無關

40 如圖所示，A 和 B 圖的 X 軸與 Y 軸的範圍相同，下列有關 A 和 B 圖的比較敘述何者錯誤？



- (A) 決定係數與相關係數都是 B 較高 (B) 相關係數是 B 較高，迴歸係數是 A 較高
(C) 迴歸係數與相關係數都是 B 較高 (D) A 的決定係數較低，但迴歸係數比 B 高