102 年公務人員普考四等 衛生行政試題

等別：四等考試
類科：衛生行政
科目：流行病學與生物統計學概要

一、研究人員比較國內兩家醫學中心醫院執行心臟移植手術後病患一年內死亡率的差異，研究結果如下表所示。研究人員發現兩家醫學中心醫院執行心臟移植手術後病患一年內死亡率都是 7%(表一)。研究人員進一步按照病患手術前罹患高血壓情形分層統計分析(表二)。
研究人員可以根據表一的統計結果，下結論認為兩家醫學中心醫院執行心臟移植手術後病患一年內死亡率沒有差異嗎？根據表二的統計數據，說明病患手術前罹患高血壓情形會如何影響比較結果？該如何處理？(10 分)

表一、兩家醫學中心醫院執行心臟移植手術後病患一年內死亡情形

	接受心臟移植手術	一年內死亡人數
醫學中心醫院甲	1000	70
醫學中心醫院乙	600	42

表二、兩家醫學中心醫院執行心臟移植手術後病患一年內死亡情形，按照病患手術前罹患高血壓情形分層

	接受心臟移植手術病患人數	一年內死亡人數
醫學中心醫院甲		
有高血壓病史	600	60
無高血壓病史	400	10
醫學中心醫院乙		
有高血壓病史	100	15
無高血壓病史	500	27

【擬答】

㈠若不考慮其他因素，單純從粗死亡率比較，兩醫學中心的一年內死亡率沒有差異。但這樣的結果並沒有考慮年齡、性別等人口學差別，以及疾病史或其他影響死亡的變因。

㈡在考慮高血壓病史的情況下，我們發現甲醫院有高血壓病史者，較容易死亡，而乙醫院有高血壓病史者，相較的死亡率較低。

甲醫院有高血壓病史者死亡率：$60 \div 600 = 0.1$

甲醫院無高血壓病史者死亡率：$10 \div 400 = 0.025$

乙醫院有高血壓病史者死亡率：$15 \div 100 = 0.15$

乙醫院無高血壓病史者死亡率：$27 \div 500 = 0.054$

所以可以採用分層的特定比率來比較，乙醫院不論是否有高血壓，皆高於甲醫院。

再者，若考慮高血壓是否是死亡的危險因子，可以分別計算其相對危險性

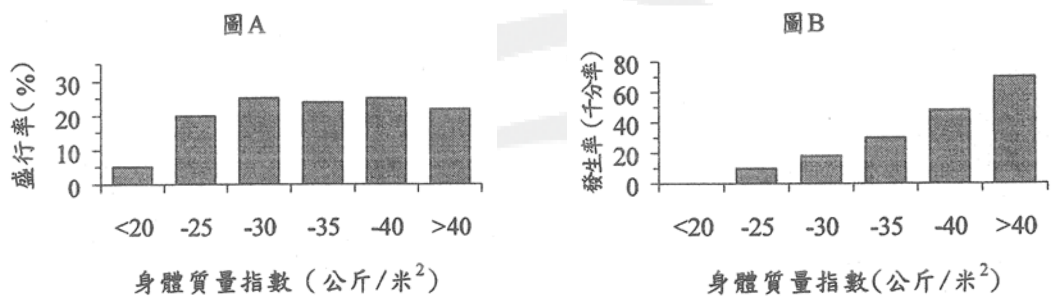
$RR_{甲} = \frac{0.1}{0.025} = 4$; $RR_{乙} = \frac{0.15}{0.054} = 2.78$

所以不論是甲醫院還是乙醫院，高血壓病史皆為一年內死亡的危險因子；而在甲醫院的高血壓危險性高過乙醫院的高血壓危險性。

二、隨著生活形態日趨禁態化及高卡路里食品的攝取，導致肥胖成為全球性健康問題。

目前最為廣泛接受脂肥胖測量指標為身體質量指數(BMI)[體重/(身高)^2，以公斤/(米)^2 為單位]。研究人員遂以身體質量指數作為肥胖測量指標，針對設定之社區民眾進行身體質量指數分布狀態與第二型糖尿病盛行率(Prevalence)(圖 A)以及身體質量指數分布狀態與第二型糖尿病發生率(Incidence)(圖 B)關聯性之分析。

請說明身體質量指數分配狀態與第二型糖尿病盛行率趨勢及發生率趨勢出現差異之原因。(20 分)



第二型糖尿病依身體質量指數類別分層之盛行率(圖 A)及發生率(圖 B)

【擬答】

由圖 B 的發生率可以得知，BMI 與罹患第二型糖尿病的風險具有劑量反應關係，即 BMI 越高，越容易有第二型糖尿病的發生。

但盛行率是觀察現存的病例，包括新發病例與舊病例，病例會在族群中隨時間慢慢累積，所以理論上也應該是 BMI 越高，第二型糖尿病累積的個案越多才對因為在穩定的情況下，

盛行率＝發生率×平均疾病期。

但從圖 A 中發現，BMI 超過 30 後，呈現一個高原的情況，並沒有直線上升，可能的原因是 BMI 超過 30 的疾病期較短。雖然第二型糖尿病是慢性病，但因為過度肥胖，除了第二型糖尿病以外的併發症更容易發生，所以可能會因為競爭死因的情況而死亡，造成現存病例的減少。這就是盛行率與發生率出現差異的主要原因。

三、國內學者針對女性多重性伴侶與否和感染 HPV(human papillomav irus)的相關性進行世代追蹤觀察研究。觀察結果如下表所示：

組別	人數	HPV 陽性人數
多重性伴侶	1000	60
單一性伴侶	4000	80

㈠在 HPV 檢驗沒有實驗誤差的情況下，女性多重性伴侶與否和感染 HPV 的相關性為何？(5 分)

㈡簡要論述 HPV 檢驗出現實驗誤差[敏感度(sensitivity)為 100%，特異度(specivicity)為 95%]，對於探討女性多重性伴侶與否和感染 HPV 相關性的影響。(10 分)

㈢解釋 HPV 檢驗特異度為 95%的意義。(5 分)

註：以下是回答下列試題可能會使用之符號(Notation)：

以 Y 代表血壓之變項(Y_0 是介入前血壓，而 Y_1 是介入後血壓)、H_0 及 H_1(代表虛無假說及對立假說)、μ_d(兩組母群體血壓差異之平均值)、σ_d^2 (母群體介入前後血壓差異之變異數)、d_i(第 i 個人前後血壓差異)、SD_d(介入前後差異之標準差)、$\bar{d}$(介入前後差異之平均值)、型一誤差(Type I error,α error (簡稱α))、n 代表樣本數、x 代表介入組別 (x=1 代表介入後，x=0 代表介入前)。(若此處未定義之符號，請以一般生物統計常識定義之)。

【擬答】

㈠因為是世代追蹤研究，所以採用相對危險來評估相關性真實相對風險 $RR=\dfrac{60/1000}{80/4000}=3$ 代表有多重性伴侶之 HPV 感染，相對單一性伴侶之 HPV 感染風險為 3 倍

㈡

	暴露組(多重)		非暴露組(單一)	
真正的分布	病例組	對照組	病例組	對照組
	60	940	80	3920
研究的分布				
病例組	60	47	80	196
對照組	0	893	0	3724
敏感度或特異度	1.0	0.95	1.0	0.95

出現實驗誤差下之相對風險 $RR=\dfrac{107/1000}{276/4000}=1.55$

多重性伴侶之 HPV 感染，相對單一性伴侶之 HPV 感染風險是 1.55 倍，故此時發生相對危險低估的情況。

㈢特異度代表無病的人，檢驗出來為陰性(無病)的機率。

故在此為實際上有感染 HPV 者，檢驗出來為感染 HPV 的機率僅有 95%

四、某縣市衛生局欲評估某降低血壓(符合常態分佈)之健康促進計畫(X=1，介入；X=0，無介入)，研究者共邀請 20 位民眾參與計畫，並測量其前後血壓值之改變。

依資料產生下列結果回答以下子題㈠及子題㈡：

介入前平均分數及標準差：$\bar{Y}_0$ (標準差,SD)=136.6(18.58)

介入後平均分數及標準差：$\bar{Y}_1$ (標準差,SD)=129.9(14.94)

$\bar{d}$=-6.7

$\Sigma_{i=1}^{k}\left(d_i-\bar{d}\right)^2=2486$ ---------------------------(1)

㈠回答下列問題：(4 分)

1. 回答上述式⑴之 k 值。

2. 計算前後差異值( )之標準差( )。

㈡若上述結果以雙尾檢定視介入前後平均血壓是否有差別，請回答下列問題：(每小題 3 分)

1. 以前述符號設立虛無及對立假說。
2. 介入前後差異平均值之變異數(Var($\bar{d}$))大小。
3. 依上述結果及在虛無假說下進行檢定，其檢定統計值大小及其自由度。
4. 計算介入前後差異平均值之 95%信賴區間。
5. 上述題 3. 判斷是否拒絕虛無假說。

㈢若將上述前後測結果重覆 500 次，請回答下列問題：(每小題 3 分，共 6 分)

1. 此 500 次樣本平均值之分佈接近何種分佈？
2. 上述題 1. 在虛無假說下之平均值及變異數。(如果原母群體前後差異之變異數(σ^2_d)為 144)

㈣若將上述前後介入視為兩組獨立樣本以簡單線性迴歸分析得到如下結果：

$$\hat{y}=\hat{\alpha}+\hat{\beta}x \qquad \text{其中 } \hat{y}\text{：血壓值, x(組別)=} \begin{cases} 1 & \text{介入後} \\ 0 & \text{介入前} \end{cases}$$

請回答下列問題。(每小題 5 分，共 25 分)

1. 利用前面介入前後所提供之平均值結果，計算 $\hat{\alpha}$ 及 $\hat{\beta}$ 。
2. 利用前面所提供介入前後之標準差計算 $\hat{\beta}$ 之標準誤。
3. 若使用兩組獨立樣本檢定，計算值。
4. 若使用一方變異數分析(One-way ANOVA)進行檢定，敘述所用檢定統計值及大小。
5. 此題 4. 分析方法相較於前面題 2. 的最大缺點是未考慮統計上的何種特性？

【擬答】

㈠ 1. $k=20$，因為是 20 組配對

2. $SD_d=\sqrt{\dfrac{\sum_{i=1}^{20}(d_i-\bar{d})^2}{20-1}}=\sqrt{\dfrac{2486}{19}}=11.4386$

㈡ 1. $H_0:\mu_d=0 \quad H_1:\mu_d\neq 0$

2. $Var(\bar{d})=\dfrac{SD_d^2}{n}=\dfrac{11.4386^2}{20}=6.5421$

3. $T^*=\dfrac{\bar{d}}{SD_d/\sqrt{n}}=\dfrac{-6.7}{11.4386/\sqrt{20}}=-2.619\sim t(19)$

服從 t 分配，自由度為 19

4. 平均差值 μ_d 之信賴區間為

$$\bar{d}\pm t_{\alpha/2}(n-1)\frac{SD_d}{\sqrt{n}}$$

$$\Rightarrow -6.7\pm 2.093\cdot\frac{11.4386}{\sqrt{20}}$$

$$\Rightarrow [-12.0534,-1.347]$$

5. $C^*:\{|T^*|>t_{0.025}(19)=2.093\}$

所以 $T^*\in C$，拒絕 H_0

有顯著的證據說介入前後平均血壓有差異。

㈢ 1. 500 筆樣本，為大樣本，可利用中央極限定理服從常態分配

2. $\bar{d} \overset{n\to\infty}{\sim} N(\mu_d, \sigma_d^2) \sim N(0, \dfrac{144}{500} = 2.88)$

㈣ 1. $x = 0 \Rightarrow \hat{Y} = \hat{\alpha} + 0 = \bar{Y}_0 \Rightarrow \hat{\alpha} = 136.6$

$x = 1 \Rightarrow \hat{Y} = \hat{\alpha} + \hat{\beta} = \bar{Y}_1 = 129.9 \Rightarrow \hat{\beta} = -6.7$

2. $\sum Y_0 = 19 \times 18.58^2 + 20 \times 136.6^2 = 379750.3116$

$\sum Y_1 = 19 \times 14.94^2 + 20 \times 129.9^2 = 341721.0684$

所以 $SS_Y = (\sum Y_0^2 + \sum Y_1^2) - 40 \times (\dfrac{136.6 + 129.9}{2})^2 = 11248.88$

$$MSE = \frac{SS_Y - \hat{\beta}^2 \times SS_X}{n-2} = \frac{11248.88 - (-6.7)^2 \times 40 \cdot 0.5^2}{40-2} = 284.21$$

$$S(\hat{\beta}) = \sqrt{\frac{MSE}{SS_X}} = \sqrt{\frac{284.21}{10}} = 5.3311$$

3. $s_p^2 = \dfrac{(n-1)SD_0^2 + (n-1)SD_1^2}{n+n-2} = \dfrac{19 \times 18.58^2 + 19 \times 14.94^2}{20+20-2} = 284.21$

$$T^* = \frac{\bar{Y}_1 - \bar{Y}_0}{\sqrt{s_p^2\left(\frac{1}{n} + \frac{1}{n}\right)}} = \frac{129.9 - 136.6}{\sqrt{284.21\left(\frac{1}{20} + \frac{1}{20}\right)}} = -1.257$$

4. $SST = \sum\sum(\bar{X}_i. - \bar{X}..)^2 = \sum n_i(\bar{X}_i. - \bar{X}..)^2$

$= 20 \cdot (136.6 - 133.25)^2 + 20 \cdot (129.9 - 133.25)^2 = 448.9$

$SSE = \sum\sum(X_{ij} - \bar{X}_i.)^2 = \sum(n_i - 1)S_i^2$

$= (20-1) \times 18.58^2 + (20-1) \times 14.94^2 = 10799.98$

$$F^* = \frac{SST / k-1}{SSE / n-k} = \frac{448.9/(2-1)}{10799.98/(40-2)} = 1.579$$

5. 事實上，題 4 的一方變異數分析的方式與題 2 的迴歸分析方法等價

即 $T^* = \left(\dfrac{\hat{\beta}}{\sqrt{\dfrac{MSE}{SS_X}}}\right) = \left(\dfrac{-6.7}{5.3311}\right)^2 = F^* = 1.579$

所以筆者懷疑本題是否是要問題 4 的一方變異數分析的方式與題(2)的相依樣本 t 檢定的差別。若是這兩種方法的差別，一方變異數分析並未考慮到相依樣本的本質含有區集效應的概念，即採用變異數分析時，應採用一因子區集設計(Randomized Block Design)。因為一因子區集設計才與相依樣本 t 檢定等價。