113 年公務人員高等考試三級考試試題

類科：衛生行政、衛生技術
科目：流行病學與生物統計學　　王瑋老師解題

一、研究者想探討某利尿劑是否有效，他從健康的民眾中隨機抽出兩組人，每組各 10 人，一組給予利尿劑，另一組給予安慰劑，然後測量其每日的尿量，所得結果如下：安慰劑組每日尿量平均值為 1180 毫升，利尿劑組每日尿量平均值為 1400 毫升，其標準差分別為 144 毫升及 245 毫升。

請問：

㈠利尿劑的利尿效果是否有達統計上顯著差異?(α=0.05)(15 分)

㈡利尿劑與安慰劑利尿效果差異的 95%信賴區間為何?(10 分)

1.《考題難易》：★☆☆☆☆
2.《破題關鍵》：獨立樣本 t 檢定與對應的信賴區間是相當基本常見的課內考題，近兩年幾乎所有高考與地特的衛生類考試皆有命，本題唯一須留意題㈠所謂之利尿效果應採用單尾檢定，如同今年普考試題檢定要採單尾，而信賴區間採雙尾進行計算。
3. 出版社書籍命中特區：王瑋，生物統計學，志光出版，頁 166～171；王瑋，生物統計學精選 500 題全解，志光出版，頁 186～195。

【擬答】

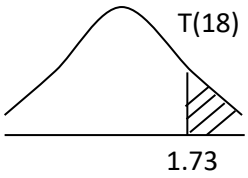
㈠考慮變異數相同下之合併變異數之 t 檢定比較平均數

$H_0:\mu_x \le \mu_y \quad H_1:\mu_x > \mu_y$

$\alpha = 0.05$

$$s_p^2 = \frac{(n-1)s_x^2 + (m-1)s_y^2}{n+m-2} = \frac{9\cdot 144^2 + 9\cdot 245^2}{10+10-2} = 40380.5$$

$$T^* = \frac{\bar{X}-\bar{Y}}{\sqrt{s_p^2\left(\frac{1}{n}+\frac{1}{m}\right)}} = \frac{1400-1180}{\sqrt{40380.5\left(\frac{1}{10}+\frac{1}{10}\right)}} = 2.448 \in C$$



$C:\{T^* > t_{18,0.95} = 1.73\}$

拒絕 H_0，有顯著的證據說

利尿劑的利尿效果達到統計上的顯著差異

㈡利尿劑與安慰劑利尿效果差異的 95%信賴區間為

$$(\bar{X}-\bar{Y}) \pm t_{18,0.975}\sqrt{s_p^2\left(\frac{1}{n}+\frac{1}{m}\right)}$$

$$\Rightarrow (1400-1180) \pm 2.10\sqrt{40380.5\left(\frac{1}{10}+\frac{1}{10}\right)}$$

$$\Rightarrow [31.279, 408.721]$$

二、某病例對照研究欲探討嚼食檳榔與口腔癌的關係，研究者找了 100 名口腔癌患者，每位患者根據年齡、性別，各配對了一名健康的民眾當對照。在此 100 組配對資料中，發現:口腔癌患者有吃檳榔，配對者有吃檳榔 共 20 對；口腔癌患者有吃檳榔，配對者無吃檳榔者共 60 對；口腔癌患者無吃檳榔，配對者有吃檳榔共 10 對；口腔患者無吃檳榔，配對者無吃檳榔者共 10 對。請問：口腔癌與嚼食檳榔有統計相關嗎?($\alpha = 0.05$) (25 分)

1.《考題難易》：★☆☆☆☆
2.《破題關鍵》：相當容易的基本題，幾乎沒有寫錯的本錢，過去也經常命題，近兩年 112 地特三等衛技、111 年普考衛生與 110 年薦任衛政皆有命題。。
3. 出版社書籍命中特區：王瑋，生物統計學，志光出版，頁 268～271；王瑋，生物統計學精選 500 題全解，志光出版，頁 338～346。

【擬答】

可將資料整理如下之配對樣本資料

口腔癌	健康對照	
病例	吃檳榔	無檳榔
吃檳榔	20	60
無吃檳榔	10	10

欲探討口腔癌與嚼食檳榔相依樣本之關聯性，採用McNemar檢定

H_0：口腔癌與嚼食檳榔無關

H_1：口腔癌與嚼食檳榔有關

$\alpha = 0.05$

$$\chi^{2^*} = \frac{(B-C)^2}{B+C} = \frac{(60-10)^2}{60+10} = 35.714 \in C$$

$$C:\{\chi^{2^*} > \chi^2_{1,0.95} = 3.84\}$$

拒絕H_0，有顯著的證據說口腔癌與嚼食檳榔有關

三、執行一個病例對照研究時，一般而言，如果病例組選自社區，則對照組也常選自社區，請說明社區對照的選取有何優點，及可能會面臨的困難。 (25 分)

1.《考題難易》：★★☆☆☆
2.《破題關鍵》：過去雖然有考過以社區為本的對照組選取方式，如 104 轉任四等衛政與 99 普考衛政，但大都考題較著重在醫院為本的對照組選取，所以本題可能對於初次應考的同學造成下筆困難。不過以本申論題的命題方式，不至於構成大失分的條件。
3. 出版社書籍命中特區：王瑋，流行病學，志光出版，頁 234～235；王瑋，流行病學精選 600 題全解，志光出版，頁 183。

【擬答】

社區對照的選取優點包括：

1. 代表性好，社區對照能夠更好地代表研究人群的基礎狀態，更貼近真實世界中的情況，因此研究結果更具代表性和外部效度。
2. 可以考慮環境因素，社區對照組可以納入環境因素對於疾病發生的影響。
3. 減少干擾，因為社區為本的對照組社經地位與居住環境相近，有助於提升研究效度。

社區對照可能的困難包括：

1. 不回應的比率較高，因為社區對照的參與度較低，這可能導致選樣偏差，進而影響研究的結果。
2. 變異性較大，不同社區或不同對象之間的差異性與一些難以控制的干擾因素，可能對研究結果產生潛在影響，重複研究的一致性可能難以保證。
3. 疾病率較低，某些疾病在社區中發生率較低的情況下，找到足夠數量的對照組可能會變得更加困難，這會影響研究結果的有效性和推廣性。

四、分層分析(stratification)是在研究分析階段時，常用來控制干擾作用的方法，請說明分層分析的意義、進行的步驟及主要的優缺點。(25 分)

1.《考題難易》：★★☆☆☆
2.《破題關鍵》：分層分析或稱為分層校正的 Mantel-Haenszel 方法過去常有解釋題與計算題型，但優缺點題型在過去未曾出現過，所以建議未來準備考試的考生，控制干擾因子的六大方法除了意義外，進行的步驟與優缺點都應該加強演練。
3. 出版社書籍命中特區：王瑋，流行病學，志光出版，頁 206～207；志聖 113 年度模擬考第二大題直接命中。

【擬答】

分層分析是根據干擾因子的觀察值分組，分別計算各組間自變項與依變項的相關強度，再根據分組人數計算加權相關強度，最常見的方法為Mantel-Haenszel method。

例如世代追蹤研究的分層校正相對危險性 $RR_{MH}=\dfrac{\sum a_i(c_i+d_i)/N_i}{\sum c_i(a_i+b_i)/N_i}$

若僅有兩個分層，公式可表達為 $RR_{MH}=\dfrac{(\frac{a_1(c_1+d_1)}{N_1})+(\frac{a_2(c_2+d_2)}{N_2})}{(\frac{c_1(a_1+b_1)}{N_1})+(\frac{c_2(a_2+b_2)}{N_2})}$

與病例對照研究的分層校正勝算比 $OR_{MH}=\dfrac{\sum(a_id_i)/N_i}{\sum(c_ib_i)/N_i}$

若僅有兩個分層，公式可表達為 $OR_{MH}=\frac{(\frac{a_1d_1}{N_1})+(\frac{a_2d_2}{N_2})}{(\frac{c_1b_1}{N_1})+(\frac{c_2b_2}{N_2})}$

分層分析的優點是採用數理計算的方式，如同加權平均，結果易於理解，簡易可行；缺點是一次僅能針對一種干擾進行校正，所以多種干擾較難同時進行分層校正，並且若分層太多造成各組樣本數太低時，計算出的結果會有不穩定的狀況。

$Z_{0.975}$ = 1.96、$Z_{0.95}$=1.645、$Z_{0.99}$ = 2.575

$t_{18,0.90}$=1.33、$t_{19,0.90}$=1.33、$t_{20,0.90}$ = 1.33

$t_{18,0.95}$=1.73、$t_{19,0.95}$=1.73、$t_{20,0.95}$ = 1.72

$t_{18,0.975}$=2.10、$t_{19,0.975}$=2.09、$t_{20,0.975}$ = 2.09

$X^2_{1,0.95}$ = 3.84、$X^2_{2,0.95}$=5.99、$X^2_{3,0.95}$ = 7.81