109 年公務人員高等考試三級考試試題

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

假說檢驗請寫出假說、計算過程及結論 $\alpha = 0.05$

一、105~109 年全國龍膽石斑養殖放養數量分別為 103.4, 143.2, 72.7, 78.1, 48.3 百萬尾。

㈠求這五年龍膽石斑養殖放養數量的平均值、中數、變異數、全距、變異係數（CV）。(10 分)

㈡求放養數量平均值的 95%信賴區間。（9 分）

㈢母群體平均值是否大於 100 百萬尾？（6 分）

【擬答】

【解題關鍵】
《考題難易》★★
《破題關鍵》先計算基本的敘述統計量，並且搭配單母體的信賴區間與假設檢定，為了連慣性，所以通常敘述統計量以樣本的形式較佳，屬常見的考題形式，如 107 高考、106 普考海洋、101 年與 100 年的原特四等皆是如此。

㈠假設龍膽石斑養殖放養數量為 X_i

平均數為 $\bar{X} = \dfrac{103.4 + 143.2 + \cdots + 48.3}{5} = 89.14$ (百萬尾)

資料重新排序：48.3, 72.7, 78.1, 103.4, 143.2

中數為 $Md = X_{(3)} = 78.1$ (百萬尾)

變異數為 $S^2 = \dfrac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1} = 1296.473$

全距為 $R = 143.2 - 48.3 = 94.9$

變異係數為 $CV = \dfrac{S}{\bar{X}} = \dfrac{\sqrt{1296.473}}{89.14} = 40.39\%$

㈡放養數量平均值之 95%信賴區間為

$\bar{X} \pm t_{0.025}(4)\dfrac{S}{\sqrt{n}}$

$\Rightarrow 89.14 \pm 2.776 \cdot \dfrac{\sqrt{1296.473}}{\sqrt{5}}$

$\Rightarrow [44.4391, 133.8409]$

㈢ $H_0 : \mu \le 100 \quad H_1 : \mu > 100$

$\alpha = 0.05$

$T^* = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}} = \dfrac{89.14 - 100}{\sqrt{1296.473}/\sqrt{5}} = -0.674 \notin C$

$C : \{T^* > t_{0.05}(4) = 2.132\}$

不拒絕 H_0，沒有顯著證據說母群體平均值大於 100 百萬尾

註：樣本平均數都沒有大於 100，檢定的結果肯定不顯著。另一方面本小題也可直接利用題(2)的信賴區間包含 100，可知母群體平均數大於 100 不會顯著。但保險起見，採用假設檢定的做法比較安心。

二、三酸甘油脂與遺傳因子 LDLR 突變可能有關，沒有此突變的 20 人三酸甘油脂平均值為 170.8（mg/dl），標準差 20.8；有突變的 10 人平均值為 250.3（mg/dl），標準差 40.4，假設三酸甘油脂為常態分佈，兩組母群體標準差相同，請問兩組三酸甘油脂平均值是否相同？（25 分）

【擬答】

【解題關鍵】

《考題難易》★

《破題關鍵》題目假設標準差相同下，故採合併變異數的 T 檢定來比較兩獨立母體平均數，屬課內超級常考的基本題，如去年 108 地特三等、108 地特四等皆有命題，可參考王瑋 生物統計學 P.5-26 與 P.5-27 頁完全相同的試題演練。

假設沒有 LDLR 突變的三酸甘油酯為 X，有 LDLR 突變的三酸甘油酯為 Y

$H_0 : \mu_x = \mu_y \quad H_1 : \mu_x \neq \mu_y$

假設 $\alpha = 0.05$

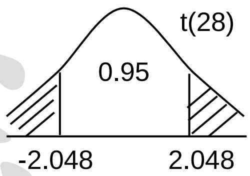
$$s_p^2 = \frac{(n-1)s_x^2 + (m-1)s_y^2}{n+m-2} = \frac{(20-1)\cdot 20.8^2 + (10-1)\cdot 40.4^2}{20+10-2} = 818.2$$

$$T^* = \frac{\overline{X} - \overline{Y}}{\sqrt{s_p^2\left(\frac{1}{n}+\frac{1}{m}\right)}} = \frac{170.8 - 250.3}{\sqrt{818.2\left(\frac{1}{20}+\frac{1}{10}\right)}} = -7.176 \in C$$

$C : \{|T^*| > t_{0.025}(28) = 2.048\}$

拒絕 H_0，有顯著的證據說

兩組三酸甘油脂平均值有差異



三、新冠肺炎確診死亡的病例中，許多原有其他病因，如心臟病、糖尿病等，下表為不同年齡有無其他病因死亡人數：

	<40	40-64	≧65
有其他病因	250	1340	2630
無其他病因	30	60	70

㈠如上表，列出原來有或無其他病因各年齡層死亡比例的期望值。（6 分）

㈡檢驗各年齡層原有其他病因與無其他病因的死亡比例是否相同？（19 分）

【擬答】

【解題關鍵】

《考題難易》★

《破題關鍵》卡方齊一性檢定屬課內基本題，只要細心不要算錯應可輕鬆拿分。可參考王瑋 生物統計學 P.7-12 至 P.7-15 頁類似的試題演練。

㈠計算各格子的期望值 $E_{ij} = \dfrac{R_i \times C_j}{n}$，並將資料整理如下，並且斜線的左上方為觀察值，右下角為期望值：

	<40	40-64	≧65	總和
有其他病因	250 / 269.77	1340 / 1348.86	2630 / 2601.37	4220

無其他病因	30 / 10.23	60 / 51.14	70 / 98.63	160
總和	280	1400	2700	4380

(二) H_0：有其他病因與無其他病因的死亡比例相同

H_1：有其他病因與無其他病因的死亡比例不相同

假設 $\alpha = 0.05$

$df = (2-1)\times(3-1) = 2$

$C:\{\chi^{2^*} > \chi^2_{0.05}(2) = 5.99\}$

$$\chi^{2^*} = \sum \frac{(O_i - E_i)^2}{E_i} = \frac{(250-269.77)^2}{269.77} + \ldots + \frac{(70-98.63)^2}{98.63} = 49.87 \in C$$

拒絕 H_0，有顯著證據說有其他病因與無其他病因的死亡比例不相同

四、海水魚類多樣性指數（Y）可能與水深（X，公尺）有關，由拖網漁船在不同深度取得 18 個網次樣本，最小平方法求得簡單直線迴歸模型如下：

變數	估計值	自由度	標準誤差
b_0	0.0004	1	0.0003
b_1	-0.92	a	0.54

(一)線性迴歸模型為何？（3 分）

(二)分別解釋迴歸係數 b0、b1 在此拖網漁船數據的意義、自由度 a 為何？（6 分）

(三)檢驗兩者是否有顯著的線性迴歸關係。（9 分）

(四)試求決定係數（coefficient of determination）並解釋其意義。（7 分）

【擬答】

【解題關鍵】

《考題難易》★★★

《破題關鍵》雖然給了三顆星，但此題的統計報表在 106 年地特三等、100 年地特三等衛技、100 年原特四等等年度皆有如出一轍的報表與問題，有作考古題的同學一定不陌生，可參考王瑋 生物統計學第八章試題演練。本題的難度在於第(4)題，只要能想到迴歸係數檢定的等價關係，此題便能迎刃而解。

(一) $\hat{Y} = 0.0004 - 0.92X$

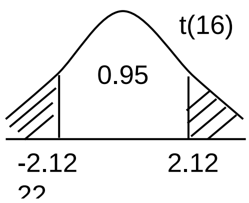
(二) b_0 代表迴歸模式的截距項，代表水深為 0 時，多樣性指數的平均值為 0.0004；b_1 代表迴歸模式的斜率項，代表每增加 1 公尺水深時，可使得多樣性指數減少 0.92 單位數。

自由度：$a = n-2 = 18-2 = 16$

(三) $H_0: \beta_1 = 0 \quad H_1: \beta_1 \neq 0$

假設 $\alpha = 0.05$

$$T^* = \frac{b_1}{S(b_1)} = \frac{-0.92}{0.54} = -1.704 \notin C$$



$C:\{|T^*| > t_{0.025}(16) = 2.120\}$

不拒絕 H_0，沒有顯著的線性迴歸關係

(四) $F^* = (n-2)\dfrac{R^2}{1-R^2} = (T^*)^2 = (-1.704)^2$

$$\Rightarrow 16 \times \frac{R^2}{1-R^2} = (-1.704)^2$$

$$\Rightarrow R^2 = 0.1536 = 15.36\%$$

代表以水深來解釋海水魚類多樣性指數的解釋力為 15.36%

