

103 年公務人員普考四等 衛生行政試題

類科：衛生行政

科目：流行病學與生物統計學概要

一、某醫療院所想進行所在鄉鎮居民之健康調查，現在考慮樣本之年齡層分布與母體之一致性，請依下列資訊計算並回答問題。

(一) 進行樣本年齡分布與母體代表性之檢測。(10 分)

(二) 請計算樣本年齡分布之中位值 (Median) 與四分位間距 (Inter-Quartile Range)。(15 分)

年齡層	樣本數	母體百分比分布
年齡 ≤ 20	15	5%
20 < 年齡 ≤ 30	25	10%
30 < 年齡 ≤ 40	35	20%
40 < 年齡 ≤ 50	50	30%
50 < 年齡 ≤ 60	32	20%
60 < 年齡 ≤ 70	20	10%
年齡 > 70	10	5%
總計	187	100%

【擬答】

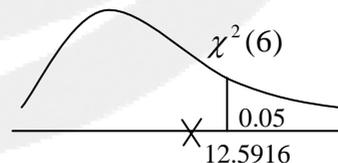
(一)

	-20	20-30	30-40	40-50	50-60	60-70	70+	總計
O_i	15	25	35	50	32	20	10	187
P_i	0.05	0.1	0.2	0.3	0.2	0.1	0.05	
E_i	9.35	18.7	37.4	56.1	37.4	18.7	9.35	

H_0 : 符合母體比例 H_1 : 不符合母體比例

$\alpha = 0.05$, $df = (7 - 1) = 6$

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} = 7.27 \notin C$$



不拒絕 H_0 ，沒有顯著證據說不符合母體比例

(二) 分組資料，中位數取第 $187 \div 2 = 93.5$ 筆，粗略觀察中位數取 45

	-20	20-30	30-40	40-50	50-60	60-70	70+
次數	15	25	35	50	32	20	10
累加	15	40	75	125	157	177	187
			Q_1	Md	Q_3		

分組資料，第一四分位數取第 $187 \div 4 = 46.75$ 筆，粗略觀察 Q_1 取 35
分組資料，第三四分位數 $187 - 46.75 = 140.25$ 筆，粗略觀察 Q_3 取 55
故四分位距為 $IQR = 55 - 35 = 20$

註：若以近似方式求解

$$Md = 40 + 10 \times \frac{93.5 - 75}{50} = 43.7$$

$$Q_1 = 30 + 10 \times \frac{46.75 - 40}{35} \doteq 31.93 ; Q_3 = 50 + 10 \times \frac{140.25 - 125}{32} \doteq 54.77$$

$$IQR \doteq 54.77 - 31.93 = 22.84$$

- 二、某減重活動發現參加者體重下降 X (單位：公斤) 與收縮壓下降 Y (單位：毫米汞柱) 有關。經初步資料統計發現以下數值： $n=50$ ；體重下降平均值 $\bar{X}=10.8$ ；體重下降標準差 $S_x = 1.44$ ；收縮壓下降平均值 $\bar{Y}=22.7$ ；收縮壓下降標準差 $S_y = 9.8$ ；體重下降與收縮壓下降共變數 $Cov_{XY} = 1.39$ 。
- (一) 請求其最小平方法之線性迴歸方程式。(15 分)
- (二) 求取體重下降 5 公斤之參加者，其平均收縮壓下降值及 95% 信賴區間。(查表值取最接近者)(5 分)
- (三) 某位參加者體重下降 10 公斤，請預測該個案之收縮壓下降值及其 95% 信賴區間。(5 分)

【擬答】

$$(一) \hat{\beta}_1 = r \cdot \frac{S_y}{S_x} = \frac{Cov_{XY}}{S_x S_y} \cdot \frac{S_y}{S_x} = \frac{1.39}{1.44^2} = 0.6703$$

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X} = 22.7 - 0.6703 \times 10.8 = 15.4608$$

$$\hat{Y} = 15.4608 + 0.6703X$$

$$(二) SSR = \hat{\beta}_1^2 \cdot SS_x = 0.6703^2 \cdot (20 - 1) \cdot 1.44^2 = 17.7018$$

$$SSTO = SS_y = (20 - 1) \cdot 9.8^2 = 1824.76$$

$$MSE = \frac{SSTO - SSR}{n - 2} = \frac{1824.76 - 17.7018}{20 - 2} = 100.3921$$

$E(Y | X = 5)$ 之 95% 信賴區間為

$$\hat{Y}_h \pm t_{\frac{\alpha}{2}}(n - 2) \sqrt{\left(\frac{1}{n} + \frac{(x_h - \bar{X})^2}{SS_x}\right) MSE}$$

$$\Rightarrow (15.4608 + 0.6703 \times 5) \pm t_{\frac{\alpha}{2}}(20-2) \sqrt{\left(\frac{1}{20} + \frac{(5-10.8)^2}{19 \cdot 1.44^2}\right) 100.3921}$$

$$\Rightarrow [-1.2012, 20.0135]$$

(三) $X = 10$ ， Y_h 之 95% 信賴區間為

$$\hat{Y}_h \pm t_{\frac{\alpha}{2}}(n-2) \sqrt{\left(1 + \frac{1}{n} + \frac{(x_h - \bar{X})^2}{SS_X}\right) MSE}$$

$$\Rightarrow (15.4608 + 0.6703 \times 10) \pm t_{\frac{\alpha}{2}}(18) \sqrt{\left(1 + \frac{1}{20} + \frac{(10-10.8)^2}{19 \cdot 1.44^2}\right) 100.3921}$$

$$\Rightarrow [0.4266, 43.9010]$$

三、流行病學進行因果關係建立時，需考慮重要干擾因子（confounder）的作用。

(一) 請定義何謂干擾因子？（6分）

(二) 在研究設定及資料分析時，有那些方法可用來處理干擾因子的作用？（15分）

(三) 未考慮重要干擾因子，可能對因果關係推論造成那些影響？（4分）

【擬答】

(一) 干擾因素指對特定危險因子對疾病的影響時，因外在因子的影響，且它可以部分或全部解釋我們所得到的因果關係時，其特性是疾病的危險因子，同時與暴露有關。

(二)

研究設計階段控制干擾因子可採用

1. 隨機分配

即研究者可運用隨機分派的技巧，將樣本隨機均勻的分組。

2. 限制法

即研究者可用限制條件來控制干擾因子，如限定年齡的範圍。

3. 配對法

即研究者可以針對干擾因子加以配對，使得干擾因子分布平均，不致影響暴露與疾病的相關性。

資料分析階段控制干擾因子可採用

1. 分層校正

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-HaensezI method。

2. 標準化

為調整不同比較團體在結構上的差異所推算出來的假想總和率，通常作的是年齡標準化，可以避免年齡組成不同所造成的影響。

3. 數學模式建立法

可以利用多變量統計分析來控制可能的干擾因子，常見的有 logistic regression 和 cox regression model 等。

(三) 因為干擾因素的影響，其可以部分或全部解釋我們所得到的因果關係，可能對於因果關係推論產生偏誤。

例如抽煙是肺癌的危險因子之一，且一般抽煙者多半同時有喝酒的習慣，因此如果研究者打算探討酒精攝取量是否會造成肺癌時，如果不控制抽煙史的分布，往往酒精攝取量與肺癌間的相關強度會受到抽煙因素的效應加以扭曲，而高估飲酒對肺癌的危險性。

四、大腸癌篩檢以糞便潛血檢查 (Fecal Occult Blood Test, FOBT) 進行，某研究單位發現其敏感度 (Sensitivity) 為 70%，而特異度 (Specificity) 為 65%，因此，以腫瘤蛋白質 X 來作為篩檢之輔助工具，發現其敏感度為 85%，而其特異度為 75%。

(一) 某社區無臨床症狀之大腸癌篩檢參與者之大腸癌盛行率為千分之二。請問以 FOBT 進行篩檢後，其陽性預測值是多少？篩檢者如何解讀該結果？(5 分)

(二) 以提高疾病被篩檢出的立場，衛生機關應以何種組合方式將 FOBT 與腫瘤蛋白質 X 應用在社區居民篩檢活動上？請計算其敏感度及特異度。(5 分)

(三) 若以 FOBT 陽性個案，再進行腫瘤蛋白質 X 工具檢測，則此整體篩檢結果將可獲得多少陽性預測值 (Positive Predictive Value)？(5 分)

【擬答】

(一)



$$PV_{+} = \frac{0.002 \times 0.7}{0.002 \times 0.7 + 0.998 \times 0.35} = 0.004$$

(二) 所有檢查同時進行，有任何一個檢查結果為陽性，就當作是有病的證據。個案必須在所有

的檢驗中都是陰性反應才會被判定為沒有生病，此為平行檢查，疾病較不可能被遺漏。
淨敏感度為兩個檢驗至少有一陽性者

$$Se = 1 - (1 - Se_1)(1 - Se_2) = 1 - 0.3 \times 0.15 = 0.955 = 95.5\%$$

淨特異度為在兩個檢驗中皆為陰性反應者

$$Sp = Sp_1 \times Sp_2 = 0.65 \times 0.75 = 0.4875 = 48.75\%$$

(三)此為序列檢查，為連續地進行檢查，是否進行下一個檢查必須前一個檢查為陽性才繼續執行。

淨敏感度為兩個檢驗中皆為陽性反應者

$$Se = Se_1 \times Se_2 = 0.7 \times 0.85 = 0.595 = 59.5\%$$

淨特異度為在兩個檢驗至少有一陰性者

$$Sp = 1 - (1 - Sp_1)(1 - Sp_2) = 1 - 0.35 \times 0.25 = 0.9125 = 91.25\%$$

所以在此疾病盛行率為0.2%，敏感度為0.595、特異度為0.9125，其陽性預測值為

$$PV+ = \frac{0.2 \times 0.595}{0.2 \times 0.595 + 0.998 \times 0.9125} = 0.1156$$

五、下表為抽菸與嚼檳榔生活習性對口腔癌之發生率分布情形，請說明依累加模式（Additive Model）及累乘模式（Multiplicative Model）進行抽菸與嚼檳榔是否有交互作用（interaction）以及交互作用型態[如協同（Synergistic）、拮抗（Antagonistic）]為何？（10分）

		嚼檳榔	
		是	否
抽菸	否	2×10^{-5}	8×10^{-5}
	是	4×10^{-5}	16×10^{-5}

【擬答】

題目很明顯有誤，正確表格應為

		嚼檳榔	
		否	是
抽煙	否	2×10^{-5}	8×10^{-5}
	是	4×10^{-5}	16×10^{-5}

並且我們將表格轉化為相對危險性

		嚼檳榔	
		否	是
抽煙	否	1	4
	是	2	8

因為 $8 > 2 + 4 - 1$ ，由 $RR_{\text{合}} > (RR_1 + RR_2) - 1$

所以檳榔及菸對口腔癌為累加的協同交互作用。

且 $8 = 2 \times 4$ ，由 $RR_{\text{合}} = (RR_1 \times RR_2)$

所以檳榔及菸對口腔癌為獨立作用，不具有累乘之交互作用。