

# 校園用水預測模式研究

張聖賢

臺北縣立三民高級中學

## 壹、研究興起

政府或學校經常宣導「節約用水」的觀念，而各種傳播媒體也配合著政府政策教導民眾如何在生活中實行節水，但是國人大多都是等到缺水警報響起的時候，才開始意識到水資源的重要性。然而，在校園中，為因應節水政策的推行，校方通常只是將校內的水槽更換成省水式的水龍頭，但我們質疑：這樣的做法僅就學校原有水槽的規劃而做的更換，並未通盤檢討配置問題，此舉真能達到節水的目的嗎？學校剛成立之初，水龍頭的配置是否已有考慮到除了能夠滿足使用者的便利性外又能達到省水的目的嗎？

為了能夠落實節約用水的政策又不失使用上的便利性，我們嘗試結合高中教材中統計學的預測方法和電腦軟體的應用尋找一個模型，用來預測及評估水龍頭的配置狀況，提供給學校一個實際且客觀的科學根據作為推行政策的參考，是我們此次研究的目標。

本人在教學過程中，有機會帶領學生，從 2005 年 9 月到 2006 年 2 月，針對本校校園內用水狀況的調查及研究。因為人力及物力有限，目前僅就部份用水機制“一般水龍頭用水”做探討，嘗試應用學生所學的知識建立起一個簡單模型後，發

現一些現象和結果，提供作為參考。另外要特別感謝我的學生陳佳凰、陳佳伶、蔡耀淳三人在這段日子裡蒐集資料與整理的辛勞，才能完成本研究。

## 貳、教材與理論介紹

### (一)、相關係數：

對兩組具有相異性質 X 與 Y 的資料數據  $x_1, x_2, \dots, x_n$  與  $y_1, y_2, \dots, y_n$ ，當  $x_1, x_2, \dots, x_n$  不盡相同，且  $y_1, y_2, \dots, y_n$  也不盡相同時，X 與 Y 之間的相關係數為

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} = \frac{S_{XY}}{S_X S_Y}$$

其中  $\bar{x} = \frac{1}{n}(x_1 + x_2 + \dots + x_n)$ ，

$\bar{y} = \frac{1}{n}(y_1 + y_2 + \dots + y_n)$ ， $S_{XY}$  為 X 和 Y

的共變異數， $S_X$  為 X 的標準差， $S_Y$  為 Y 的標準差。相關程度可區分為：

- (1)  $r=1$ ，稱 X 與 Y 為完全正相關。
- (2)  $r=-1$ ，稱 X 與 Y 為完全負相關。
- (3)  $r=0$ ，稱為零相關。
- (4)  $0 < |r| < 0.3$ ，表示 X 與 Y 為低度相關。
- (5)  $0.3 \leq |r| < 0.7$ ，表示 X 與 Y 為中度相關。
- (6)  $0.7 \leq |r| < 1$ ，表示 X 與 Y 為高度相關。

**(二)、簡單線性迴歸 (Simple Linear Regression) 與最小平方法：**

給定 X 與 Y 兩種抽樣數據  $x_1, x_2, \dots, x_n$  與  $y_1, y_2, \dots, y_n$ 。我們想要探討 X 與 Y 兩種變量是否有線性的關係，因此欲尋找一條直線  $y = a + bx$  來代表這個線性關係。

要如何評估這條直線的代表性？我們是以殘差的大小作為評估的準則。每一筆資料的殘差( $e_i$ )定義為第 i 筆資料的觀

測值與擬合值  $\hat{y}_i = a + bx_i$  的差，即

$$e_i = y_i - \hat{y}_i, \quad i = 1, \dots, n$$

而這一條直線  $y = a + bx$  擬合此組資料的整體表現是以殘差平方和(SSE)

$$SSE = \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

當做評估的標準，當 SSE 愈小，直線的擬合程度愈好。

根據上面的想法，我們要求二實數 a 與 b，使

$$SSE = \sum_{i=1}^n (y_i - a - bx_i)^2 \dots\dots\dots(式 1)$$

的值為最小。這樣找出來的直線  $y = a + bx$  稱為 Y 對 X 的簡單線性迴歸直線，而這種方法稱為最小平方法。利用 Y 對 X 的迴歸線，我們可以由 x 值預估相應的 y 值，或由 y 值預估相應的 x 值。

經由數學運算結果我們可以知道：欲使 SSE 的值最小，則 a 與 b 需滿足下列方程組：(此方程組可將式 1 用偏微分算出。)

$$\begin{cases} a + b\bar{x} - \bar{y} = 0 \\ b - \frac{S_{XY}}{S_{XX}} = 0 \end{cases}$$

解得  $a = \bar{y} - \frac{S_{XY}}{S_{XX}}\bar{x}$ ， $b = \frac{S_{XY}}{S_{XX}}$ 。

其中  $S_{XY}, S_{XX}$  與  $S_{YY}$  為：

$$S_{XY} = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) = \frac{1}{n-1} (\sum_{i=1}^n x_i y_i - n\bar{x}\bar{y})$$

$$S_{XX} = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 = \frac{1}{n-1} (\sum_{i=1}^n x_i^2 - n\bar{x}^2)$$

$$S_{YY} = \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2 = \frac{1}{n-1} (\sum_{i=1}^n y_i^2 - n\bar{y}^2)$$

由上面的計算結果可以得知此組資料的最佳迴歸直線為：

$$\begin{aligned} y &= (\bar{y} - \frac{S_{XY}}{S_{XX}}\bar{x}) + \frac{S_{XY}}{S_{XX}}x \\ &= \bar{y} + r \cdot \frac{S_Y}{S_X}(x - \bar{x}) \end{aligned}$$

其中， $S_X$  為 X 的標準差， $S_Y$  為 Y 的標準差，r 為相關係數。

**(三)、判定係數：**

如果想要知道一組資料其迴歸式對該組資料的擬合程度，直接想到的是計算此組資料的殘差平方和(SSE)，

$$SSE = \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

當 SSE 愈小，直線迴歸式的擬合程度愈好。但是 SSE 會受到單位的影響，因此通常以判定係數  $R^2$  來作為判斷的指標，

$$R^2 = 1 - \frac{SSE}{SSTO}, \text{ 其中 } SSTO = \sum_{i=1}^n (y_i - \bar{y})^2$$

因為直線迴歸式是由最小平方方法求得的所有直線最能擬合此資料的直線，也就是殘差平方和最小的，所以  $SSE < SSTO$ 。

又  $SSTO - SSE$  表示利用迴歸直線做擬合時降低的殘差平方和，稱為迴歸差平方和  $SSR$ ，即  $SSR = SSTO - SSE$ ，所以

$$SSR = \sum_{i=1}^n (y_i - \bar{y})^2 - \sum_{i=1}^n (y_i - \hat{y}_i)^2 = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2$$

因此

$$R^2 = 1 - \frac{SSE}{SSTO} = \frac{SSTO - SSE}{SSTO} = \frac{SSR}{SSTO}$$

當  $R^2$  愈大，此迴歸線愈有解釋此組資料的能力。

#### (四)、多元線性迴歸 (Multiple Linear Regression) :

此處我們僅以二變數為例，介紹二元線性迴歸，其模式為：

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2}, \quad i = 1, 2, \dots, n$$

在上式中，我們引進了兩個自變數： $X_{i1}$  和  $X_{i2}$ ，而參數  $\beta_1$  可解釋成當  $X_{i1}$  增加一單位時，導致應變數  $Y_i$  改變的比例；同理  $\beta_2$  為當  $X_{i2}$  增加一單位時導致  $Y_i$  改變的比例。那麼  $\beta_0$ 、 $\beta_1$  和  $\beta_2$  要如何估計呢？

假設存在著  $n$  組樣本，第一組樣本為  $(X_{11}, X_{12}, Y_1)$ ，第二組樣本為  $(X_{21}, X_{22}, Y_2)$ ，...，依此類推。接下來我們得根據這些樣本觀察值對未知參數  $\beta_0$ 、 $\beta_1$  和  $\beta_2$  進行估計。這三迴歸參數的估計量仍可用最小平方方法求得。令殘差平方和  $SSE$  為

$$SSE = \sum_{i=1}^n [Y_i - (\beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2})]^2$$

我們要找出可以使  $SSE$  值為極小值的  $\beta_0$ 、 $\beta_1$  和  $\beta_2$ 。符合這項條件的  $\beta_0$ 、 $\beta_1$  和  $\beta_2$  分別用符號  $b_0$ 、 $b_1$  和  $b_2$  來表示。將上式分別對  $b_0$ 、 $b_1$  和  $b_2$  作偏微分，並令為零，得到以下列三條方程式：

$$\sum Y_i = nb_0 + b_1 \sum X_{i1} + b_2 \sum X_{i2}$$

$$\sum X_{i1} Y_i = b_0 \sum X_{i1} + b_1 \sum X_{i1}^2 + b_2 \sum X_{i1} X_{i2}$$

$$\sum X_{i2} Y_i = b_0 \sum X_{i2} + b_1 \sum X_{i1} X_{i2} + b_2 \sum X_{i2}^2$$

利用聯立方程式的技巧，可解得  $b_0$ 、 $b_1$  和  $b_2$ 。如此便可找到了參數式  $\beta_0$ 、 $\beta_1$  和  $\beta_2$  的估計量： $b_0$ 、 $b_1$  和  $b_2$ 。如此便可得到樣本迴歸函模式： $\hat{Y} = b_0 + b_1 X_1 + b_2 X_2$ ，我們用此式來估計二元線性迴歸模式。

#### (五)、抽樣誤差與樣本數的決定：

習慣上我們都以 95% 的信賴係數做為一般抽樣設計的常模，因此公式  $2\sigma/\sqrt{n}$  就成為我們決定樣本數和誤差大小的依據。

有一個比較快速的估計方式是使用百分比，也就是將以前的平均數現在變成百分比，亦即從 0 到 1 之間的一個數字。樣本百分比的標準差則跟著這個百分比變化，但是絕對不會超過  $0.5/\sqrt{n}$ 。為了保險起見，我們就用  $0.5/\sqrt{n}$  來代入，換句話說，原來的公式  $2\sigma/\sqrt{n}$ ，現在變成  $2 \times 0.5/\sqrt{n} = 1/\sqrt{n}$ ，這是估計的最大抽樣誤

差。例如樣本數為 400 時，抽樣誤差為  $\pm 1/20 = \pm 0.05$ 。樣本數為 900 時，抽樣誤差是  $\pm 0.033$ 。我們可以發現樣本數在 1000 到 1600 時最划得來。若再往上加，抽樣工作負擔加重很多，但抽樣誤差卻減少得很有限，並不經濟。所以 1000 到 1600 是最常見到的樣本數。

### 參、研究過程及方法

(一) 調查全校的水龍頭分佈和總數，以及每棟大樓的教室配置。完成調查後，為方便資料的蒐集，於是將所有的水龍頭逐一編號，以便記錄。

(二) 我們探討的是「正常上課時間」用水狀況，因此僅利用課間的時段來作資料的蒐集，將觀察的時間區分成下面十個時段來做觀測。

1 : 7 : 20 → 7 : 30	6 : 12 : 00 → 12 : 10
2 : 8 : 00 → 8 : 10	7 : 12 : 55 → 13 : 00
3 : 9 : 00 → 9 : 10	8 : 13 : 50 → 14 : 00
4 : 10 : 00 → 10 : 10	9 : 14 : 50 → 15 : 10
5 : 11 : 00 → 11 : 10	10 : 16 : 00 → 16 : 10

(三) 依據教室分佈的比例決定每棟樓應抽出的樣本數，並要求每個時段抽取等量的樣本下，利用「簡單隨機抽樣法」作為抽取樣本的準則，來決定所選取的樣本。本次研究的過程我們分別在秋初(月均溫 25°C)蒐集了共 50 個樣本數，冬天(月均溫 17.7°C) 蒐集了共 50 個樣本數。

(四) 利用攝影機先將使用者的用水情況拍攝下來後再進行資料的整理和記

錄，這種方法可以降低資料收集時所產生的人為誤差。

(五) 為了能獲得用水量的資訊，我們利用量杯和碼錶測量了校園中兩種水龍頭類型的出水量得知：「省水型」每秒出水量 35C.C.，而「不省水型」每秒出水量 360C.C.。

(六) 利用 Microsoft Excel 與 Equation Grapher 等軟體整理和分析資料。

(七) 參考高中教材的相關係數與直線相關的課程，分「秋初」和「冬」兩個不同時節，分別依「時段與使用人數」、「時段與使用時間」、「時段與用水量」三種情況計算所對應的相關係數及簡單迴歸直線方程式。

(八) 利用高中教材所介紹的方法以原始的數據作簡單線性迴歸時，如果僅將時段以 1、2、...、10 代入計算，其預測的結果會出現在某一個時段所預測出來的用水量為負值的不合理情況。為了解決這個問題，我們考慮將每一個時段量化，因此引用「權數 (weight)」的概念作為量化的手段，方法是將每一個時段的資料按照其所佔全部資料的比例當作該時段的權數 ( $0 < \text{weight} < 1$ )，然後再進行簡單線性迴歸分析，最後利用此模型預測相對時段單一洗手槽之使用人數、使用時間及用水量。同時，為檢驗引進權數的概念是否可以提升簡單線性迴歸的準確性，也做了多元線性迴歸模型的計算，檢視與簡單線性

迴歸所做的預測結果是否一致。

(九) 爲了探討影響用水的分佈、用水習慣和目的，我們設計了一份調查問卷，在考慮各年級的班級數及教室的分佈比例後，決定每棟樓應該抽取的樣本數，利用「分層抽樣」作爲抽取樣本的準則，決定所選取的樣本。此次問卷調查的過程，我們發出 1000 張

問卷，實際收回 934 張。抽樣是在 95% 的信賴度下，最大抽樣誤差爲 $\pm 3.3\%$ 。

## 肆、調查及研究結果：

### (一) 資料整理：

1、依據各時段將原始數據中，每一洗手槽使用人數、使用時間和用水量整理如下：

時段與使用人數統計表(秋初)

單位：人次

時段 洗手槽	1	2	3	4	5	6	7	8	9	10
	7:20 7:30	8:00 8:10	9:00 9:10	10:00 10:10	11:00 11:10	12:00 12:10	12:55 13:00	13:50 14:00	14:50 15:10	16:00 16:10
1				3	2		0	2	30	10
2			3	11	13	11	9	1	15	3
3										1
4	7	0	1	12	4				20	11
5			0	2						5
6							0	5	11	
7	5	2	25	1	2	5	1	11	45	
8	5	0			1	6	6	3		
9	7	3	11			12				
10	8	1				2				
合計 總計	32	6	40	29	22	36	16	22	121	30
	354									

時段與使用人數統計表(冬)

單位：人次

時段 洗手槽	1	2	3	4	5	6	7	8	9	10
	7:20 7:30	8:00 8:10	9:00 9:10	10:00 10:10	11:00 11:10	12:00 12:10	12:55 13:00	13:50 14:00	14:50 15:10	16:00 16:10
1	5	0	3	6	7	0	0	0	13	4
2		2	0	6	15	0	7	8	8	0
3		0	0	0		7	0	0	8	2
4			24	0			0	0	10	
5	2	3		16	5	5	1	2	10	2
6	3	2			5	0				1
7	4				2					
8	4				2					
合計 總計	18	7	27	28	36	12	8	10	49	9
	204									

時段與使用時間統計表(秋初)

單位：Sec.

時段 洗手槽	1	2	3	4	5	6	7	8	9	10
	7:20 7:30	8:00 8:10	9:00 9:10	10:00 10:10	11:00 11:10	12:00 12:10	12:55 13:00	13:50 14:00	14:50 15:10	16:00 16:10
1				11	3		0	7	137	23
2			6	80	130	101	35	3	75	16
3										3
4	25	0	7	103	28				56	42
5			0	20						18
6							0	44	49	
7	16	28	135	1	9	35	5	63	288	
8	22	0			5	16	25	8		
9	25	10	55			57				
10	34	1				5				
合計	122	39	203	215	175	214	65	125	605	102
總計	1865									

時段與使用時間統計表(冬)

單位：Sec..

時段 洗手槽	1	2	3	4	5	6	7	8	9	10
	7:20 7:30	8:00 8:10	9:00 9:10	10:00 10:10	11:00 11:10	12:00 12:10	12:55 13:00	13:50 14:00	14:50 15:10	16:00 16:10
1	13	0	6	9	27	0	0	0	68	6
2		8	0	37	46	0	25	20	99	0
3		0	0	0		26	0	0	77	3
4			80	0			0	0	53	
5	7	6		71	12	9	2	3	40	3
6	2	2			8	0				1
7	6				4					
8	12				12					
合計	40	16	86	117	109	35	27	23	337	13
總計	803									

時段與用水量統計表(秋初)

單位：C.C.

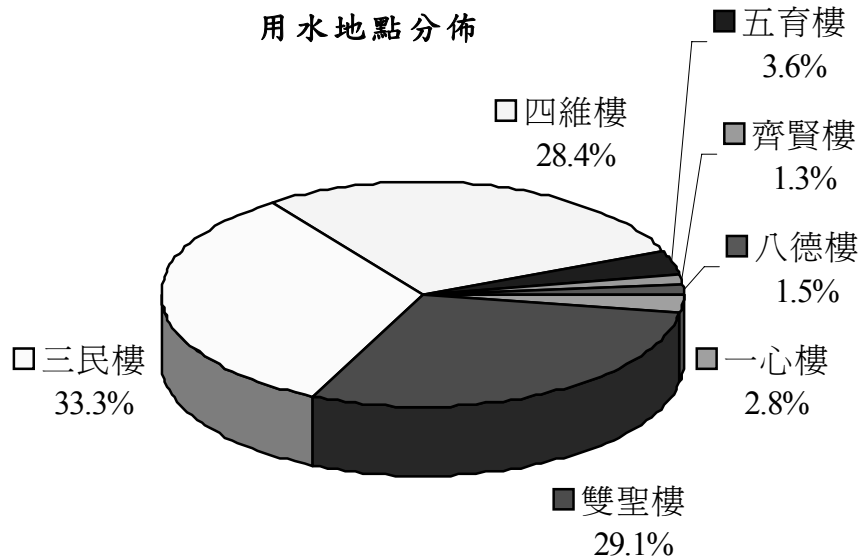
時段 洗手槽	1	2	3	4	5	6	7	8	9	10
	7:20 7:30	8:00 8:10	9:00 9:10	10:00 10:10	11:00 11:10	12:00 12:10	12:55 13:00	13:50 14:00	14:50 15:10	16:00 16:10
1				3960	1080		0	2520	53920	8420
2			2160	7025	46800	32460	12600	1080	25375	5760
3										1080
4	9000	0	2520	37080	10080				18535	15120
5			0	7200						6480
6							0	15840	17640	
7	5760	10080	48600	360	3240	12600	1800	22680	88410	
8	7920	0			1800	5760	8350	1905		
9	9000	3600	19800			20520				
10	12240	360				1800				
合計	43920	14040	73080	55625	63000	73140	22750	44025	203880	36860
總計	630320									

時段與用水量統計表(冬)

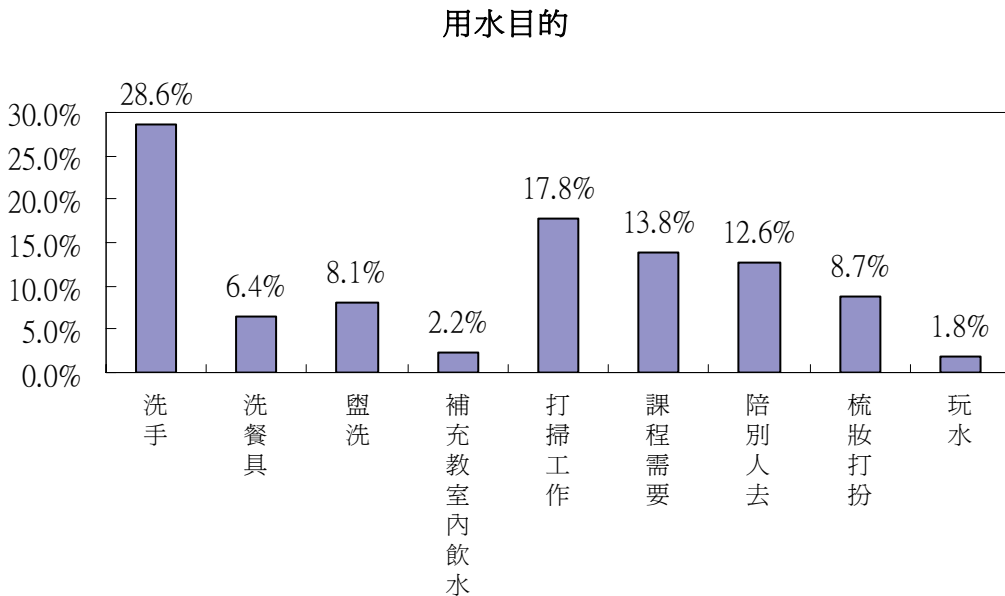
單位：C.C.

時段 洗手槽	1	2	3	4	5	6	7	8	9	10
	7:20 7:30	8:00 8:10	9:00 9:10	10:00 10:10	11:00 11:10	12:00 12:10	12:55 13:00	13:50 14:00	14:50 15:10	16:00 16:10
1	3055	0	2160	3240	9720	0	0	0	22212	2160
2		2880	0	4220	16560	0	9000	7200	35640	0
3		0	0	0		9360	0	0	27720	1080
4			28800	0			0	0	19080	
5	2520	2160		25560	4320	3240	720	1080	13428	1080
6	720	720			2880	0				360
7	2160				1440					
8	3600				4320					
合計	12055	5760	30960	33020	39240	12600	9720	8280	118080	4680
總計	274395									

2、用水地點抽樣的分佈統計：學校上課用教室絕大多數集中在雙聖、三民、四維等四棟樓，其餘均為專科教室或辦公室。

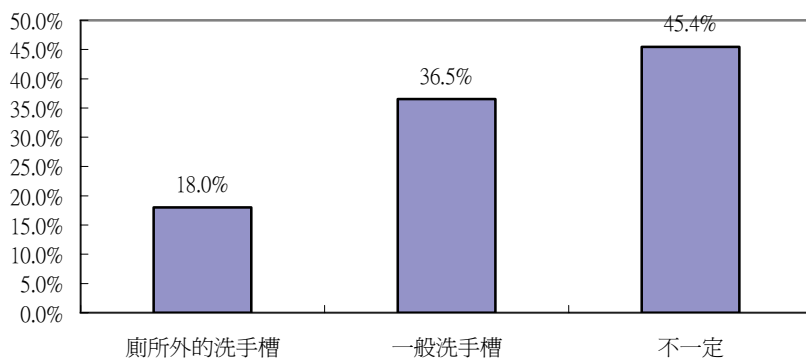


3、問卷調查結果統計：有效樣本數為 934 人，在 95%的信賴度下，最大抽樣誤差 $\pm 3.3\%$ 。

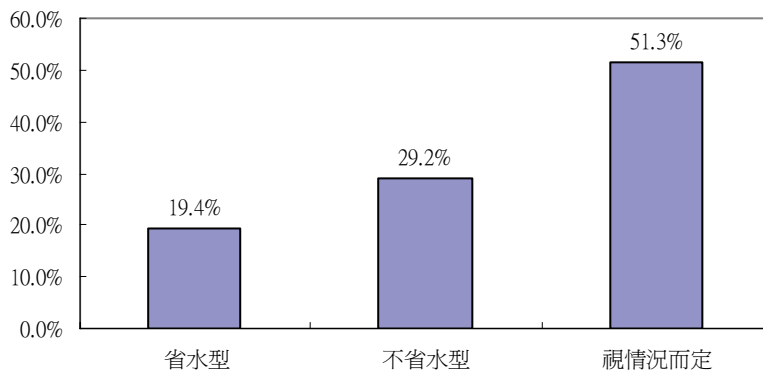




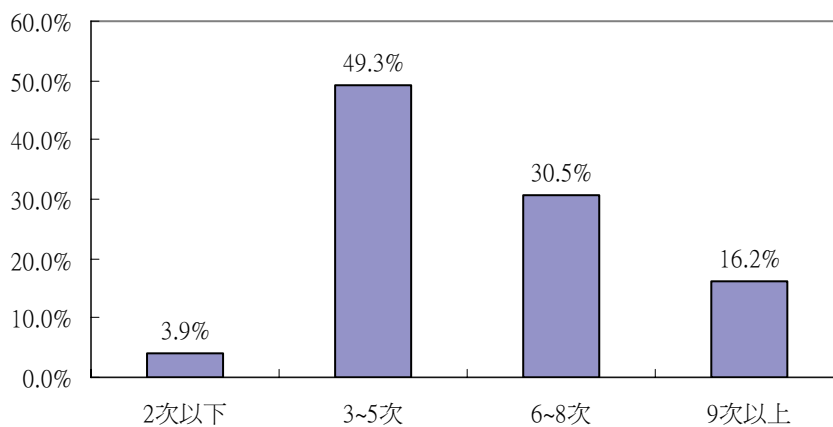
洗手槽的選擇



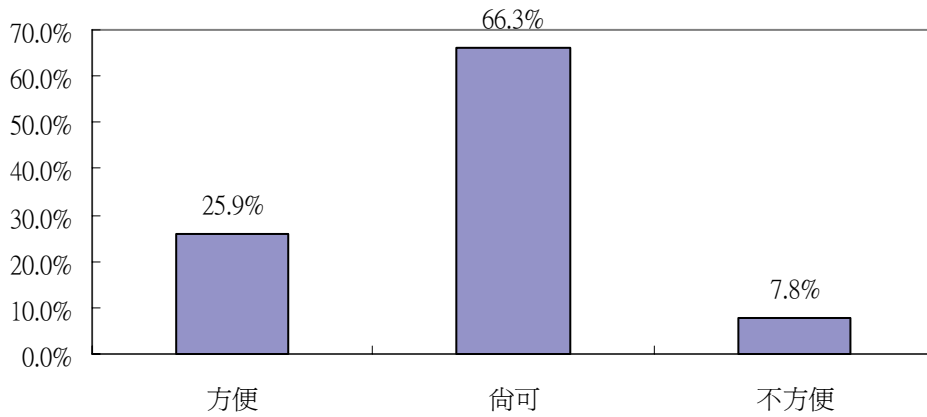
水龍頭類型的選擇



一天用水次數



便利性



(二) 資料分析：

1、各時段之使用人數、使用時間及用水量的權數(weight)：爲了能夠利用簡單線性迴歸模型合理預測使用人數、使用時間及用水量或是藉由多元線性迴歸模型做用水量的檢驗，我們計算了在不同變因的情況下各時段的權數。在此僅以「時段與使用人數(秋初)」之「時段 1」爲例說明計算方法，其餘依此類推，整理如下表：

$$\begin{aligned}
 \text{時段 1 的權數} &= \text{該時段次數和} \div \text{總次數} \\
 &= 32 \div 354 \\
 &\approx 0.0904
 \end{aligned}$$

時段	(秋初)			平均
	使用人數	使用時間	用水量	
1	0.0904	0.0654	0.0706	0.0755
2	0.0169	0.0209	0.0226	0.0201
3	0.113	0.1088	0.1175	0.1131
4	0.0819	0.1153	0.0895	0.0956
5	0.06	0.0938	0.1013	0.085
6	0.1017	0.1147	0.1177	0.1114
7	0.0452	0.0349	0.0366	0.0389
8	0.0642	0.067	0.0708	0.0673
9	0.342	0.3245	0.3164	0.3276
10	0.0847	0.0547	0.057	0.0655

(冬)				
時段	使用人數	使用時間	用水量	平均
1	0.0882	0.0498	0.0439	0.0606
2	0.0343	0.0199	0.021	0.0251
3	0.1324	0.1071	0.1128	0.1174
4	0.1373	0.1457	0.1203	0.1344
5	0.1765	0.1357	0.143	0.1517
6	0.0588	0.0436	0.0459	0.0494
7	0.0392	0.0336	0.0354	0.0361
8	0.049	0.0286	0.0302	0.0359
9	0.2402	0.4198	0.4304	0.3636
10	0.0441	0.0162	0.0171	0.0258

2、在簡單線性迴歸模型下各變量間的相關係數及判定係數：

相關係數和判定係數	(秋初)		(冬)	
	相關係數	判定係數	相關係數	判定係數
時段&使用人數	0.7200	51.84%	0.5294	28.03%
時段&使用時間	0.5938	35.26%	0.7512	56.43%
時段&用水量	0.6009	36.11%	0.7412	54.94%
使用人數&使用時間	0.9364	87.68%	0.8701	75.71%
使用人數&用水量	0.9327	86.99%	0.8710	75.86%
使用時間&用水量	0.9804	96.12%	0.9888	97.77%

(三) 預測模式與結果：

前面曾提及我們研究的目的是希望利用高中教材中所介紹的「簡單線性迴歸」尋找預測模型，而為了檢驗簡單線性迴歸模型來預測不同時段之用水量的可靠性，因此同時利用了「多元線性迴歸模型」做用水量的預測來比較之間的差異。但為了進一步計算多元線性迴歸模型的預測結果，也利用簡單線性迴歸計算了「時段與使用人數」、「時段與使用時間」的預測結果。下面的結果是以原始數據為根據計算出單一洗手槽之預測值。

1、簡單線性迴歸模式下預測使用人數：

(秋初)			(冬)		
$Y=0.00294+70.8*\text{時段 weight}$			$Y=0.13484+39.1*\text{時段 weight}$		
時段	時段 weight	預測人數(人次)	時段	時段 weight	預測人數(人次)
1	0.0904	6	1	0.0882	4
2	0.0169	1	2	0.0343	1
3	0.113	8	3	0.1324	5
4	0.0819	6	4	0.1373	6
5	0.06	4	5	0.1765	7
6	0.1017	7	6	0.0588	2
7	0.0452	3	7	0.0392	2
8	0.0642	5	8	0.049	2
9	0.342	24	9	0.2402	10
10	0.0847	6	10	0.0441	2

2、簡單線性迴歸模式下預測使用時間：

(秋初)			(冬)		
$Y=0.00311+373.0*\text{時段 weight}$			$Y=0.00884+159.6*\text{時段 weight}$		
時段	時段 weight	預測使用時間(sec)	時段	時段 weight	預測使用時間(sec)
1	0.0654	24	1	0.0498	8
2	0.0209	8	2	0.0199	3
3	0.1088	41	3	0.1071	17
4	0.1153	43	4	0.1457	23
5	0.0938	35	5	0.1357	22
6	0.1147	43	6	0.0436	7
7	0.0349	13	7	0.0336	5
8	0.067	25	8	0.0286	5
9	0.3245	121	9	0.4198	67
10	0.0547	20	10	0.0162	3

## 3、簡單線性迴歸模式下預測用水量：

(秋初)			(冬)		
$Y = -288.39929 + 129437.27 * \text{時段 weight}$			$Y = 5.25513 + 54512.46 * \text{時段 weight}$		
時段	時段	預測用水量(cc)	時段	時段 weight	預測用水量(cc)
1	0.0706	8850	1	0.0439	2398
2	0.0226	2637	2	0.021	1150
3	0.1175	14920	3	0.1128	6154
4	0.0895	11296	4	0.1203	6563
5	0.1013	12824	5	0.143	7801
6	0.1177	14946	6	0.0459	2507
7	0.0366	4449	7	0.0354	1935
8	0.0708	8876	8	0.0302	1652
9	0.3164	40666	9	0.4304	23467
10	0.057	7090	10	0.0171	937

4、多元線性迴歸模式下的預測結果：在多元線性迴歸模型中，若要求算 Y 值時，其各變數必須落在原來樣本的範圍內，經比較所預測出的使用人數、使用時間和原始數據之後，發現這些預測值均落在原來樣本的範圍內。又由於校園內各處所收集到的數據落差甚大，難以斷定以哪一組的數據作為基準值較為合理，因此我們嘗試以簡單線性迴歸所預測的使用人數和使用時間來預測用水量。為了謹慎起見，在相同的模型下，我們也將所有的原始數據做了用水量的推測，下頁表中各時段的預測用水量和原始數據所作的預測用水量平均值的誤差大部分在±1%範圍內，接近我

們所做的推測。為建立後續研究的準則，我們擬以簡單線性迴歸預測的使用人數和時間作為此處計算多元線性迴歸的基準，而不失其準確性。

5、預測每月用水度數與實際用水度數之比較：51 頁表中分別利用「簡單線性迴歸模型」和「多元線性迴歸模型」在不同時節時的預測用水量做整理比較，並據此推估使用度數。至於使用度數如何計算？因為我們收集到的數據為單一洗手槽各水龍頭使用狀況的和，而每一洗手槽有 1~3 個水龍頭不等，經調查全校總共有 299 個水龍頭，所以平均一個洗手槽約有 2.52 個水龍頭。然後將每日預測用水量(單一洗手

槽)除以 2.52 再乘上 299，即可得每日每一時段全校的預估總用水量，求其總和，即得每日全校的預估總用水量。又以每週上課 5 日、每月 4.3 週計，將每日全校之預估總用水量乘上 5 再乘以 4.3，可計算出每月全校在上課時間內

的預估總用水量(升)。因為自來水 1 立方公尺(1000 公升)的水量 = 1 度，所以將每月全校預估總用水量除以 1000 後，就可算出預測的使用度數。(若干數值因求取近似值之緣故，有些許誤差。)

(1) Simple Linear Regression :

(秋初)

(冬)

$$Y=185.58-222.65*X1+244.05*X2+287.27*X3$$

$$Y=-283.43+915.17*X1+87.64*X2+331.36*X3$$

時段	時段平均權數	人數	時間 (sec)	預測用水量 (c.c)
1	0.0755	6	24	8740
2	0.0201	1	8	2714
3	0.1131	8	41	13773
4	0.0956	6	43	13936
5	0.085	4	35	11256
6	0.1114	7	43	14210
7	0.0389	3	13	4699
8	0.0673	5	25	8461
9	0.3276	24	121	40794
10	0.0655	6	20	7497

時段	時段平均權數	人數	時間 (sec)	預測用水量 (c.c)
1	0.0606	4	8	2723
2	0.0251	1	3	924
3	0.1174	5	17	5956
4	0.1344	6	23	8030
5	0.1517	7	22	7651
6	0.0494	2	7	2284
7	0.0361	2	5	1676
8	0.0359	2	5	1445
9	0.3636	10	67	23088
10	0.0258	2	3	763

Correlation 0.9812

Correlation 0.9891

Y : 用水量      X2 : 人數      X3 : 時間  
X1 : 時段平均權數

Y : 用水量      X2 : 人數      X3 : 時間  
X1 : 時段平均權數

(2) Multiple Linear Regression :

(秋初)			(冬)		
時段	預測用水量 (升)/日	預測全校總用 水量(升)/日	時段	預測用水量 (升)/日	預測全校總用 水量(升)/日
1	8.85	1050.04	1	2.40	284.57
2	2.64	312.87	2	1.15	136.45
3	14.92	1770.33	3	6.15	730.21
4	11.30	1340.31	4	6.56	778.72
5	12.82	1521.53	5	7.80	925.54
6	14.95	1773.40	6	2.51	297.50
7	4.45	527.88	7	1.93	229.59
8	8.88	1053.12	8	1.65	195.96
9	40.67	4825.00	9	23.47	2784.43
10	7.09	841.18	10	0.94	111.23
總計	126.55	15015.65	總計	54.57	6474.18

預測使用度數	實際使用度數	比例
323	4171	7.7%

預測使用度數	實際使用度數	比例
139	4035	3.4%

(秋初)					(冬)				
時段	預測人 數/日	預測用水 間(秒)/日	預測用水 量(升)/日	預測全校總用 水量(升)/日	時段	預測人 數/日	預測用水 間(秒)/日	預測用水 量(升)/日	預測全校總用 水量(升)/日
1	6	24	8.74	1037.02	1	4	8	2.72	323.05
2	1	8	2.71	322.04	2	1	3	0.92	109.66
3	8	41	13.77	1634.13	3	5	17	5.96	706.74
4	6	43	13.94	1653.47	4	6	23	8.03	952.79
5	4	35	11.26	1335.51	5	7	22	7.65	907.86
6	7	43	14.21	1686.02	6	2	7	2.28	270.98
7	3	13	4.70	557.55	7	2	5	1.68	198.81
8	5	25	8.46	1003.86	8	2	5	1.44	171.41
9	24	121	40.79	4840.28	9	10	67	23.09	2739.43
10	6	20	7.50	889.56	10	2	3	0.76	90.50
總計			126.08	14959.46	總計			54.54	6471.23

預測使用度數	實際使用度數	比例
322	4171	7.7%

預測使用度數	實際使用度數	比例
139	4035	3.4%

經過上面的比較，以簡單線性迴歸來預測時段與用水量的模型其相關係數在秋初時節為 0.6009、在冬的時節為 0.7412，呈中度相關。但引進權數之後，用簡單線性迴歸推估整個月的用水度數和以多元線性迴歸所做出來的結果，一致性卻相當的高，可推測到使用人數、時間、用水量間為高度正相關，這似乎可以說明僅利用簡單線性迴歸和適當的權數即可作為用水量的預測，而毋須用較複雜的多元線性迴歸模型來計算，此結果可作為後續研究之參考。

## 伍、討論

- (一) 為提高調查效率，本研究採樣的樣本僅限於設置在校園走廊上的水槽，其他如廁所、實驗室或專科教室內所使用的水槽並未考慮。而且我們利用攝影機來收集資料，這種方式可以降低人為觀測上的誤差。
- (二) 在原始數據中，有些時段某些地點的記錄為零，而有些時段(如：打掃時間)使用人數又異常的多，顯示這些地點的水龍頭的設置可做檢討。推測發生這種現象的原因與教室的配置、課程的安排和特定執行的任務有相當程度的關係。
- (三) 從預測的結果中看出：洗手槽的用水佔全校用水的比例並不高。在秋初的時節僅佔全部用水的 7.7%，在冬的時節佔 3.4%。因此校方推行節水政策時，應該要有更全面性的討論才行。
- (四) 從問卷調查中整理出「用水地點分

布」的結果來看，習慣用水地點集中分佈在雙聖、三民、四維樓，而這些樓層為一般教室，使用人數共佔了 90.8%，其餘為專科教室。結果顯示用水狀況與教室的分佈有密切的關係。

- (五) 問卷調查中顯示：一般人若有不同類型的水龍頭可以選擇時，「視情況而定」的人佔了 51.3%，選擇「不省水型」的人佔了 29.2%，選擇「省水型」的佔 19.4%。在調查過程中發現，多數人仍會優先選擇不省水型的水龍頭使用。
- (六) 問卷調查中也顯示：學生在學校用水的主要目的以洗手 28.6%居冠，打掃工作 17.8%居次。有趣的是，有些人原本沒有用水意願，但是當陪伴別人去時，竟也打開水龍頭用起水來，這佔了 12.6%，排名第四。還有，純粹玩水也佔了 1.8%。這些數據顯示出青少年用水觀念的多樣性。
- (七) 從問卷調查中發現：每人每天平均用水最高是 3~5 次，佔 49.3%；其次是 6~8 次，佔 30.5%；再來是 9 次以上，佔 16.2%；最後是 2 次以下次，佔 3.9%。從數據可以看出：雖然正常用水的人佔絕大多數，但也有一天使用 9 次以上的人，這似乎述說部份青少年用水的無目的性。
- (八) 從問卷中可以了解：大多數學生對於水龍頭的配置並沒有強烈的認知意見。目前校方安排水龍頭的方式，幾



乎可以全面滿足用水需求，只有少部分的人(7.8%)會覺得不方便。或許可以進一步去探究不方便的原因，作為改善的參考。

- (九) 此次研究的目的是希望能找出一個高中生可以應用的預測模型來推估用水狀況。經過半年的資料收集和分折，我們利用高中教材中的「簡單線性迴歸」模型，加上在「時段」以權數量化以後，找到了可行的預測公式，此與較複雜的多元線性迴歸模型所做的預測結果比較，有相當的一致性，因此一般高中生可以在現有的學習基礎下作後續的研究。但是要如何去找到一個可以全面預測全校用水的模型，目前在觀測技術上、變因控制上仍有許多困難要克服，但此次的研究，應可以提供未來更為複雜相關研究之參考。

## 陸、未來展望

做此研究之前，曾實地去了解各機關或學校單位對水費預算的提報後發現：大多單位均以往年的數據加加減減之後作為下一年度提報預算的基準，鮮少看到檢討實際用水狀況的調查或改進的可能性。當初在決定執行這項計畫時，是希望透過統計學的工具，在條件的配合下，能夠找出一種預測用水情況的方法，讓學校或其他

機關能有一個客觀依據來落實節水政策，並有效控制預算的提報、支出與監控，也藉此過程來教育學生如何去做一份科學的研究。

本次研究雖僅掌握學校用水不到 10% 的狀況，但在研究過程中也慢慢累積出一些方向和心得，縱然仍有 90% 以上的狀況有待釐清，但只要克服其中的困難，慢慢便能將整體的實際狀況以客觀的結果呈現，屆時便能對用水狀況做全面的預測、管控與改進。

## 柒、參考資料

- 余文卿主編(民 94)：高中數學教材甲。台北縣：龍騰文化。
- 李虎雄、陳昭地等(民 93)：高中數學第四冊。台中市：康熙圖書。
- 林福來等(民 93)：高中數學第四冊。台南市：南一書局。
- 徐宗昌、薛玉共(民 74)：「統計學」。聯邦書局，。
- 詹世煌譯 (Hoel、Port、Stone 原著)(民 74)：「統計導論」。台北市：曉圓出版社。
- 吳柏林(民 88)：「現代統計學」，二版。台北市：五南圖書。
- 陳義彥等(民 90)：「民意調查」。台北市：五南圖書。
- Robert V.Hogg&Elliot A.Tanis(1983)：「Probability and Statistical Inference」，2<sup>nd</sup> edition。Macmillan Publishing Co.,Inc.，New York。
- 台灣自來水公司網站：  
[http://www.water.gov.tw/04service/ser\\_c\\_main.asp](http://www.water.gov.tw/04service/ser_c_main.asp)

## 【附件】

我們是本校高中部的學生，目前正在做「校園用水狀況」的研究，爲了能了解校園內用水與供水狀況的適合性，亟需你的協助，提供寶貴資料作爲參考。爲了使調查出來能儘可能符合實際的情形，請大家在回答時務必依照個人的習慣勾選出最適當的選項，或於畫線處填明。本資料僅提供研究之用並絕對保密，敬請惠予協助，謝謝各位！

### 一. 基本資料

1. 國中 高中                      2. 班級: \_\_\_\_\_年 \_\_\_\_\_班                      3. 性別: 男 女

### 二. 你經常在學校裡面的何處用水?(至多選三項)

【說明：以”面對教室爲參考方向”填「左」、「中左」、「中」、「中右」、「右」】

- 一心樓\_\_\_\_\_樓    左 中左 中 中右 右 的洗手槽  
 雙聖樓\_\_\_\_\_樓    左 中左 中 中右 右 的洗手槽  
 三民樓\_\_\_\_\_樓    左 中左 中 中右 右 的洗手槽  
 四維樓\_\_\_\_\_樓    左 中左 中 中右 右 的洗手槽  
 五育樓\_\_\_\_\_樓    左 中左 中 中右 右 的洗手槽  
 齊賢樓\_\_\_\_\_樓    左 中左 中 中右 右 的洗手槽  
 八德樓(科學館)\_\_\_\_\_樓    左 中左 中 中右 右 的洗手槽

### 三. 下面是調查同學們在學校時用水的習慣和目的，請仔細回答：

#### 1. (1)在學校時，通常在什麼情況下會去用水?(可複選)

- 洗手 洗餐具(含洗杯子) 盥洗 補充教室內的飲用水 打掃工作  
課程需要(家政.美術.實驗……) 陪別人去 梳妝打扮 玩水  
其他\_\_\_\_\_

(2)其中最主要的是哪一項? \_\_\_\_\_

#### 2. 在學校時，如果你想要用水，你會選擇哪裡?(單選)

- 廁所外的洗手槽    一般的洗手槽    不一定

#### 3. 一個洗手槽中可能有不同的水龍頭類型，一般你會選擇什麼類型來用?(單選)

- 省水型                      不省水型                      視情況而定

#### 4. 你在學校生活一整天下來，每天大概會用水幾次?(單選)

- 2次以下                      3~5次                      6~8次                      9次以上

#### 5. 你覺得現在學校洗手槽的配置情形，對你而言，使用上方不方便?

- 方便                      尚可                      不方便