

變異數分析與實驗設計

--綠豆成長觀察實驗

統計三 106304020 林子勛

統計三 106304025 宋承翰

統計三 106304027 許智超

統計三 106304042 方嘉呈

統計三 106304050 張韶恩

目錄

壹、	研究動機	2
貳、	研究目的	2
參、	實驗設計	2
1.	treatment design :	2
2.	Experiment design :	2
肆、	實驗分析	3
伍、	結論	5
陸、	附錄	6
1.	圖	6
▲	圖 一	6
▲	圖 二	6
▲	圖 三	6
▲	圖 四	7
▲	圖 五	7
▲	圖 六	7
▲	圖 七	7

壹、 研究動機

豆芽菜因為不易腐壞，菜價較低穩定，再加上種植不用土壤，只需要水就可以在室內生長，是種植最簡單的蔬菜。在颱風來襲時，豆芽菜是非常重要的替代蔬菜。我們想知道除了純水以外，像是酸性的水溶液，又或者是含鹽分的水溶液，能不能使豆芽菜成長茁壯，甚至效果能夠比純水還好。

貳、 研究目的

一般來說，種植豆芽菜的菜農都使用純水來種植豆芽菜，然而，近年來有關酸雨和土壤鹽化對植物的影響受到大幅的報導與研究，於是我們想說假如不是使用純水來種植豆芽菜，而是使用酸性的水溶液，或者水分有鹽的成分，能否使豆芽菜成長得更有效率。

參、 實驗設計

1. treatment design：

為了做此實驗，我們設計了兩大變因：兩種不同酸鹼度和三種不同鹽度的水。

酸鹼度的部分，選取純水(ph7)、醋水溶液(ph5.9)。酸雨的 ph 值是 5.5，而 5.9 是我們比較好取得的酸鹼度。原本醋的 ph 值為 2.9，以醋:水=1:1000 的比例，調出 ph 值 5.9 的醋水溶液。

鹽度部分，選取 0%、0.05%、0.1%的鹽水溶液。根據農委會的報告綠豆的耐鹽度為 0.1%，故我們選擇鹽度為 0%、0.1%以及處於這兩種濃度中間的 0.05%來做實驗。

2. Experiment design：

準備 120 顆綠豆，在飲水機溫水中浸泡一晚。再來將綠豆分成 6 組，分別放置在 6 張衛生紙上並以 6 種不同的溶液澆灌，每天澆 10ml，重複 7 天。第八天測量綠豆芽的長度，觀察用哪種溶液澆灌的綠豆芽長最高，怎樣的酸鹼度和鹽度最適合綠豆芽生長。

肆、實驗分析

首先，關於數據的部分，我們是以 csv 檔的形式呈現。檔案中第一行就是我們種的綠豆的長度，也是我們的反應變數 y 。第二行是第一個變因：酸鹼度，分為純水與醋兩項。第三行是第二個變因：鹽度，分為 0%、0.05%、0.1%。

在完成數據輸入後，我們將此 csv 檔讀進 R 中，然後使用 `tapply` 求出在不同變因底下，綠豆長度的平均值與標準差。其結果見圖一。其中，1 號為純水鹽度 0%，2 號為純水鹽度 0.05%，3 號為純水鹽度 0.1%，4 號為酸性水溶液鹽度 0%，5 號為酸性水溶液鹽度 0.05%，6 號為酸性水溶液鹽度 0.1 %

	1.	2.	3.	4.	5.	6.
平均	20.15	16.75	15.75	15.70	19.90	15.55
標準差	7.97	8.19	2.07	5.50	6.73	4.9

接下來，我們製作了盒鬚圖以及變因的交互作用圖（詳見圖二、圖三）。從盒鬚圖（圖二）中，可以看出在醋溶液中，隨著鹽度的增加，綠豆長度的平均值呈現先增後減的情況；相較之下，在純水溶液中，隨著鹽度的增加，綠豆長度的平均值呈現先減後增的情形。在交互作用圖（圖三）中，可以明顯地看出兩個變因有很明顯的交互作用。

再來，為了檢測變異數的同質性，我們進行了一個虛無假設為變異數同質，對立假設為變異數不同值，且顯著水準為 0.05 的 `levene test`，其結果見圖四。結果中 $p\text{-value}=0.1987$ ，大於我們訂定的顯著水準 0.05，不拒絕虛無假設，即代表變異數同質性假設成立。

我們做出了 ANOVA Table（圖五），從表中可以看到酸鹼度、鹽度，以及酸鹼度與鹽度的交互作用他們的 $p\text{-value}$ ，分別為 0.4893、0.0338、0.0024。我們所設定的顯著水準為 0.05，而鹽度以及鹼度與鹽度的交互作用這兩項的 $p\text{-value}$ 都小於顯著水準，因此意味著鹽度以及鹼度與鹽度的交互作用項為顯著結果。

酸鹼值	鹽度	交互作用
P=0.489	P=0.034	P=0.002

ANOVA Table 做完後，我們進行了正交比較(orthogonal contrasts)的分析，其結果見圖六。結果為：鹽度的一次項、酸鹼度和鹽度的二次項的交互作用這兩項顯著。

最後，我們利用正交比較(orthogonal contrasts)分析的結果幫助求出正交迴歸多項式模型。首先，我們算出各行的平均分別為 20.15（純水，鹽度 0%）、

16.75 (純水，鹽度 0.05%)、15.75 (純水，鹽度 0.1%)、15.70 (酸性水溶液，鹽度 0%)、19.90 (酸性水溶液，鹽度 0.05%)、15.55 (酸性水溶液，鹽度 0.1%)，為了求出模型中各係數 (Effect) 之值，我們利用公式

$\frac{\sum P_{cij} \bar{y}_{ij}}{\sum P_{cij}^2}$ 得到常數項、酸鹼值(PH)、鹽度一次項(SI)、鹽度二次項(Sq)、酸鹼值與鹽度一次項之交互作用(PHSI)、酸鹼值與鹽度二次項之交互作用(PHSq)的係數，分別是 17.225, -0.025, -0.5, 1, 1.3125, -0.6625，並且在正交比較分析，在顯著水準為 0.05 時，藉由 p-value 得出只有 SI 及 PHSq 顯著，因此模型中只保留其二並刪除不顯著之變數。

酸鹼值(PH)	鹽度一次項(SI)	鹽度二次項(Sq)	(PHSI)	(PHSq)
P=0.48927	P=0.01050	P=0.64648	P=0.12634	P=0.00168

	常數項	酸鹼值	鹽度一次項	鹽度二次項	PHSI	PHSq
$\frac{\sum P_{cij} \bar{y}_{ij}}{\sum P_{cij}^2}$	17.225	-0.025	-0.5	1	1.3125	-0.6625

我們得到的正交迴歸模型如下。

$$\hat{Y} = 17.225 - 0.5 * SI - 0.6625 * PHSq$$

關於常態性假設的部分，我們畫了 qqplot (圖七)。該圖中的直線是常態直線圖，當 qqplot 散佈圖在此直線附近時，代表樣本資料來自常態分佈。在圖三中，可以看到資料點皆落在常態直線圖附近，可以推論出我們的迴歸模型服從常態假設。

接下來我們為了檢測各項變因是否對綠豆的生長有影響，製作了一個虛無假設為該變因對綠豆的生長沒有影響，對立假設為變因對綠豆的生長有影響的 anova table，結果如圖五。表中可由 p-value 是否大於我們設定的顯著水準 0.05，得知鹽度和鹽度與酸鹼值的交互作用項為顯著，即不否認該變因對綠豆生長沒有影響，而酸鹼度這項變因的檢定結果不顯著，可以說酸鹼值並不影響綠豆的生長。

伍、 結論

從上述分析我們可以看出「鹽度」與「鹽度跟酸鹼值的交互作用」皆為顯著，「酸鹼值」則是不顯著。

也就是說，鹽度對綠豆芽的生長有顯著影響，酸鹼值雖然本身不顯著，但是由於交互作用顯著，代表酸鹼對綠豆芽的生長還是會有影響。

而根據我們的實驗結果，「鹽度 0% 的純水」澆灌出來的綠豆芽生長的最好。根據迴歸模型，固定其他變數不變的情況下，鹽度每增加一個水準，綠豆芽的高度會減少 $(0.5 + 0.6625 * \text{PH 值})$ 毫米。同樣固定其他變數的時候，PH 每增加一個水準，綠豆芽的高度就會減少 $(0.6625 * \text{鹽度})$ 毫米。可以看到，鹽度跟酸鹼對綠豆生長都會有影響，而且這個影響是負面的。

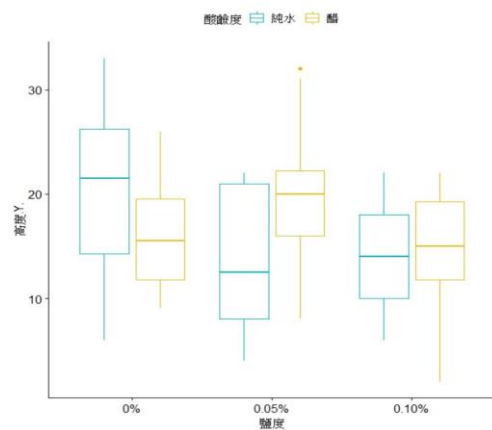
由這個結果我們可以大膽推論，土壤鹽化跟酸雨對綠豆的生長都有負面影響。在種綠豆時應避免這兩個變因，盡量使用鹽度為 0%，PH 值為 7 的純水，這樣綠豆才能長得最好。

陸、 附錄

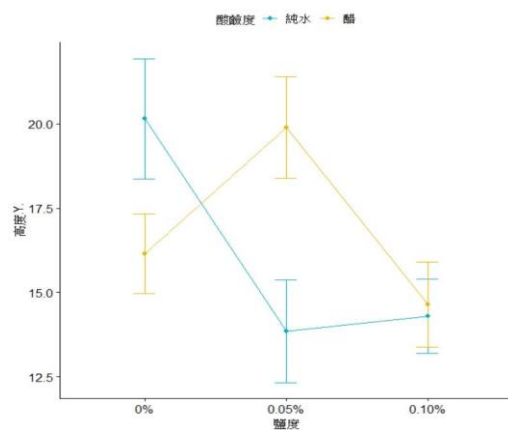
1. 圖

```
> tapply(ex22[,2], ex22[,1],mean)
 1      2      3      4      5      6
20.15 16.75 15.75 15.70 19.90 15.55
>
> tapply(ex22[,2], ex22[,1],sd)
 1      2      3      4      5      6
7.968854 8.187764 2.074279 5.497368 6.734827 4.904080
```

▲ 圖 一



▲ 圖 二



▲ 圖 三

```
Levene's Test for Homogeneity of Variance (center = median)
      Df F value Pr(>F)
group  5  1.4894 0.1987
114
```

▲ 圖 四

```
> DATA.anova <- aov(高度.Y. ~ 酸鹼度 * 鹽度, data = DATA)
> summary(DATA.anova)
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
酸鹼度	1	19	19.2	0.481	0.48927
鹽度	2	279	139.3	3.491	0.03377 *
酸鹼度:鹽度	2	508	254.0	6.367	0.00239 **
Residuals	114	4548	39.9		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

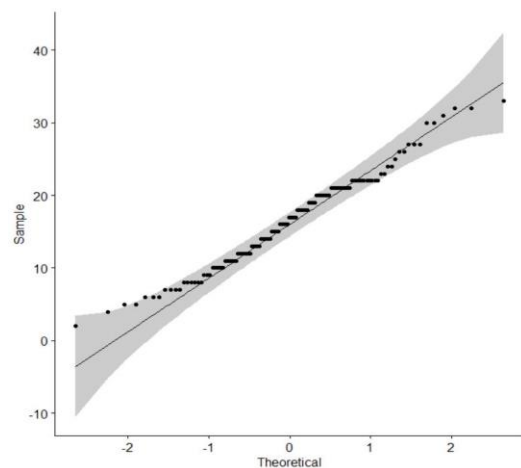
▲ 圖 五

```
> DATA.orc <- aov(高度.Y. ~ PH + S1 + Sq + PHS1 + PHSq, data = DATA)
> summary(DATA.orc)
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
PH	1	19	19.2	0.481	0.48927
S1	1	270	270.1	6.770	0.01050 *
Sq	1	8	8.4	0.211	0.64648
PHS1	1	95	94.6	2.371	0.12634
PHSq	1	413	413.4	10.363	0.00168 **
Residuals	114	4548	39.9		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

▲ 圖 六



▲ 圖 七