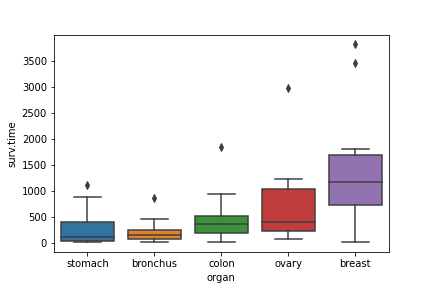
HW2 RE6091020 汪玄同

6.5

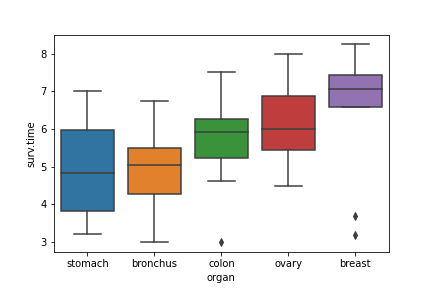
資料由Cameron及Pauling收集，資料中紀錄不同器官得癌症後，接受抗壞血酸治療後的存活時間，下表為資料變數說明。

|  |  |
| --- | --- |
| Surv.time | 患者治療後的存活時間 |
| organ | 癌症器官 |

存活時間轉換前的盒形圖



存活時間經過log轉換後的盒形圖



轉換前資料每一個類別內的資料較為分散變異數較大，轉換後的資料較為集中變異數較少，假如資料的每一個類別對於資料的值有指數關係，log轉換可以消除這層指數關係，使作檢定或配飾回歸線時可以擬合數據的分布。

對經過轉換的資料做變異數分析(ANOVA)其模型如下:

同質性檢定:

虛無假設、對立假設如下

得到的檢定統計量與P-value為

在0.05的顯著水準下檢定結果不拒絕虛無假設。資料當作每個類別標準差皆相同，每個類別沒有隨機效應的影響，使用固定效應的ANOVA模型檢定

統計量與P-value為

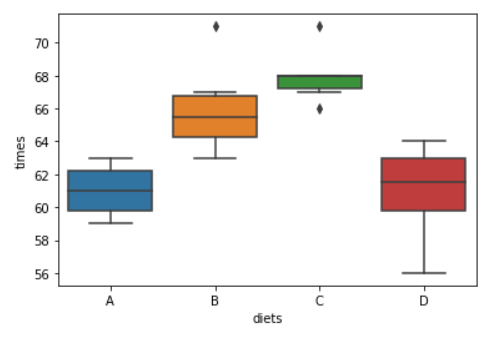
在顯著水準0.05下由於p-value < 0.05，拒絕虛無假設，**存在一組**

最後檢查殘差是否為常態

檢定統計量為 0.9650， P-value為0.06699，由於 P-value 小於0.05不拒絕虛無假設，殘差當作常態分佈。

6.9

資料由Box於1978年收集，資料以四種不同的飲食方式對應到從動物體內抽取血液的凝固時間。比較不同飲食對血液凝固的影響，下圖為各組飲食方式的盒形圖



從合型圖中可以看出沒個飲食習慣對應的血液凝固時間範圍不同。

使用變異數分析ANOVA比較每組血液平均凝固時間是否相同，模型如下

先檢定同質性，檢查組間是否變異數相同，假設如下

檢定統計量與P-value為

結果不拒絕虛無假設，組間的變異數視為相同，沒有隨機效應的影響。

使用固定效應的變異數分析ANOVA檢定，假設如下

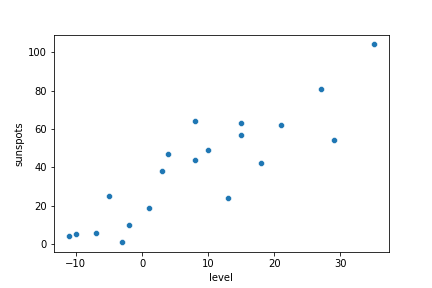
檢定統計量為 0.9783，P-value為0.8628，不拒絕虛無假設。

使用tukey多重比較檢定各組每一對之間期望值是否相同，

|  |  |  |  |
| --- | --- | --- | --- |
|  | 區間下界 | 區間上界 | P-value |
| A對B | 0.7243 | 9.2756 | 0.0183 |
| A對C | 2.7243 | 11.2756 | 0.001 |
| A對D | -4.0562 | 4.0562 | 0.9 |
| B對C | -1.8242 | 5.8242 | 0.4774 |
| B對D | -8.5772 | -1.4227 | 0.0044 |
| C對D | -10.5772 | -3.4227 | 0.001 |

結果A對D，B對C，沒有顯著差異。

8.2

資料收集太陽黑子數對標準水位的水平，將兩資料做回歸模型。

從圖來看水位對太陽黑子數呈現正相關。

建立回歸模型如下

檢定太陽黑子數是否對水位有影響，也就是檢定Beta是否為0，假設如下

計算得到統計量與P-value分別為 7.82, 0.0000003

拒絕虛無假設，太陽黑子樹對水位有影響，Beta不為0

使用Shapiro-Wilk檢定殘差服從常態，假設如下

統計量與P-value為0.956, 0.479

不拒絕殘差服從常態。

使用Durbin Watson檢查殘差是否獨立，假設如下

得到統計檢定量與P-value為 1.707, P-value為0.352

不拒絕殘差獨立假設

使用Breusch Pagan檢查殘差是否符合同質性，假設如下

統計量與P-value為0.063, 0.802

不拒絕殘差具同質性假設。