## 卡方練習題:

題目一

對於某一卡通，學生(分成：小學生、中學生、大學生三類)的喜好程度(分成：喜歡、普通、不喜歡三種)如下表所示：

|  |  |  |  |
| --- | --- | --- | --- |
|  | 小學生 | 中學生 | 大學生 |
| 喜歡 | 54 | 66 | 72 |
| 普通 | 62 | 78 | 84 |
| 不喜歡 | 49 | 65 | 70 |

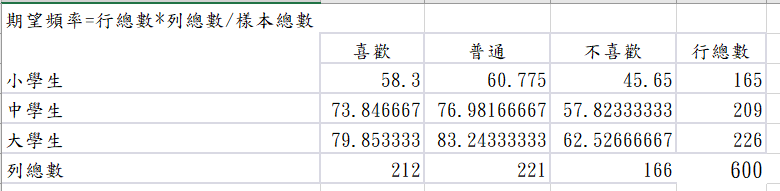
試以顯著水準0.05，檢定學生分類與喜歡程度是否獨立?

設定假設：

H0：學生分類與喜歡程度為獨立

H1：學生分類與喜歡程度非獨立

計算期望頻率：



計算卡方值 = Σ [(觀察頻率 – 期望頻率)² / 期望頻率]：

[(54-58.3)²/ 58.3 + (62-60.775)²/ 60.775 + (49-45.65)²/ 45.65 + (66-73.84)²/ 73.84 + (78-76.98)²/ 76.98 +(65-57.82)²/ 57.82 + (72-79.85)²/ 79.85 + (84-83.24)²/ 83.24 +(70-62.52)²/ 62.52 ] =3.99

自由度：(4-1)x(4-1)=9，依照卡方分配表(α=0.05)取得臨界值 9.488

* 由於卡方值小於臨界值(3.99 < 9.488)，故學生分類與喜歡程度為獨立。

題目二

某工廠為研究時間不同之工作者所生產的產品不良率是否相同，於是進行隨機抽樣得資料如下表所示：

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | 早班 | 晚班 | 夜班 | 總和 |
| 不良品 | 42 | 50 | 65 | 157 |
| 良品 | 913 | 900 | 880 | 2693 |
| 總和 | 955 | 950 | 945 | 2850 |

試檢定(α=0.05)，產品不良率與工作時間之不同是否有差異？

設定檢定：

H0：產品不良率與工作時間並無差異

H1：產品不良率與工作時間存在差異

計算機率：

1. 早班不良率：42/955 ≈ 0.044，良率：913/955 ≈ 0.956
2. 晚班不良率：50/950 ≈ 0.053，良率：900/950 ≈ 0.947
3. 夜班不良率：65/945 ≈ 0.069，良率：880/945 ≈ 0.931

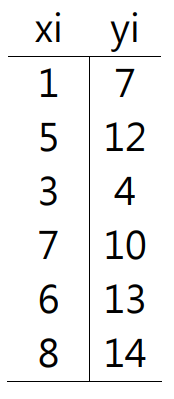
計算期望值：

1. 早班期望不良數：(955/2850) x 42 ≈ 14.08，期望良數：(955/2850) x 913 ≈ 899.92
2. 晚班期望不良數：(950/2850) x 50 ≈ 16.67，期望良數：(950/2850) x 900 ≈ 900
3. 夜班期望不良數：(945/2850) x 65 ≈ 21.25，期望良數：(945/2850) x 880 ≈ 858.75

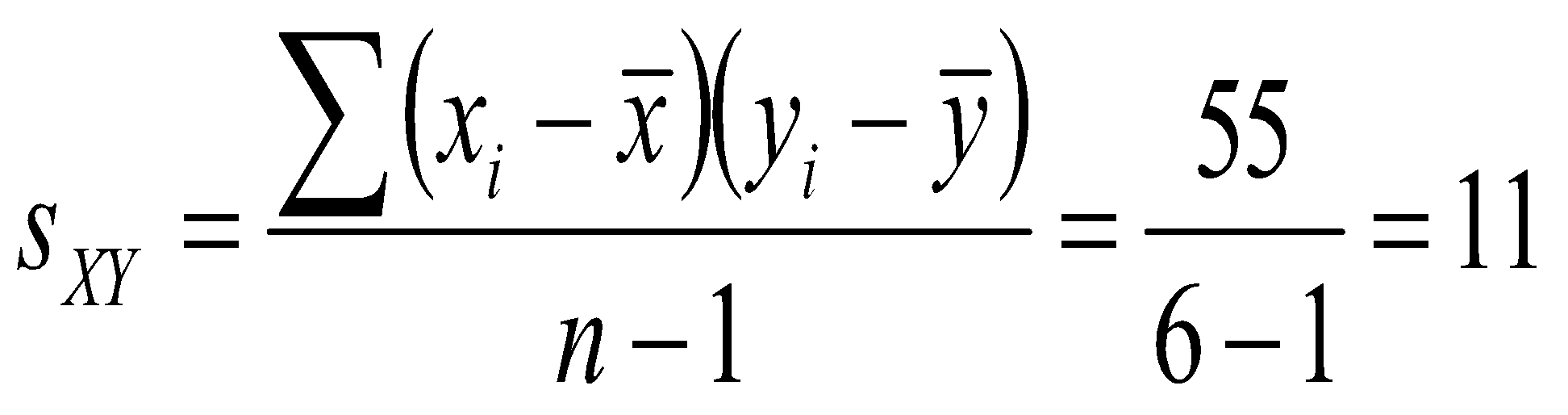
卡方檢定的統計量 = (42 - 14.08)^2/14.08 + (913 - 899.92)^2/899.92 + (50 - 16.67)^2/16.67 + (900 - 900)^2/900 + (65 - 21.25)^2/21.25 + (880 - 858.75)^2/858.75 ≈ 12.39

自由度：3-1 =2，依照卡方分配表(α=0.05)取得臨界值 5.99

* 由於統計量大於臨界值(12.39>5.99)，故拒絕H0假設，產品不良率與工作時間存在差異。



**請跟據右表計算 sample covariance (樣本共變異數)：**

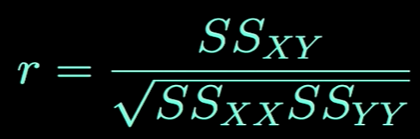


Xi 平均值：(1 + 5 + 3 + 7 + 6 + 8) / 6 = 5

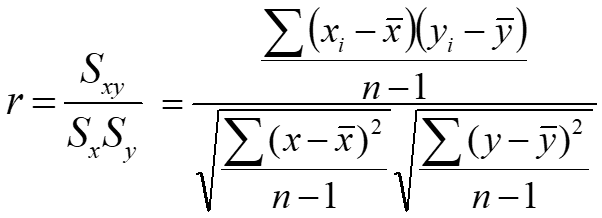
Yi 平均值：(7 + 12 + 4 + 10 + 13 + 14) / 6 = 10

* 樣本共變異數(sample covariance) = [(1-5)(7-10) + (5-5)(12-10) + (3-5)(4-10) + (7-5)(10-10) + (6-5)(13-10) + (8-5)(14-10)] / 5 = 7.8

**請跟據前一題計算 sample Pearson correlation coefficient（樣本相關係數）**

****

**(SS = Sum of Squared)**

****

**(S = Standard Deviation)**

Xi 標準差：sqrt([(1-5)^2 + (5-5)^2 + (3-5)^2 + (7-5)^2 + (6-5)^2 + (8-5)^2] / 5) = 2.6

Yi 標準差：sqrt([(7-10)^2 + (12-10)^2 + (4-10)^2 + (10-10)^2 + (13-10)^2 + (14-10)^2] / 5) = 3.84

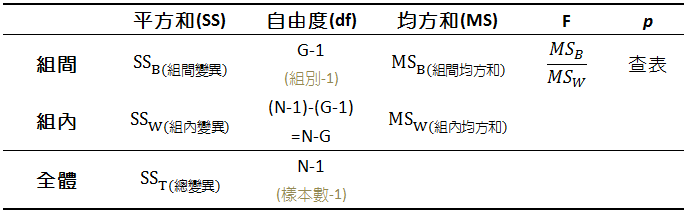
* **r = 7.8 / (2.6 \* 3.84) = 0.78**

**請跟據前一題計算 R squared 並解釋其意義。**

* **由於 r 為 0.78 代表 X 與 Y 存在高度正相關**

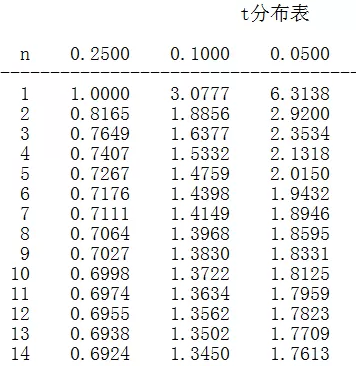
**請根據前一題數據，再加上下面這筆資料，完成ANOVA表，並解釋其結論。**

Zi = 9,5,2,1,4,3



**請跟據前一題數據，使用Student’s T-test 並解釋其結論。**

**附圖:**





**F分配表(α=0.05)**

