統計方法應用---卡方檢定　　　應數博 96　林良靖

拿到一筆資料，最想知道的就是這筆資料來自何種分佈；而卡方檢定是可以廣泛的檢定出所得資料是否服從某種 分佈的其中一種方法 。在這篇文章將 以 Berkson (1966) 從美國國家標準局所獲得的資料來做分析；從一種金屬鋂 (Am, 原子序 95)中所釋放出的α放射粒子，在連續的放射中紀錄 10,220 次，觀測值的平均放射速率為：$\dfrac{(\text{total \#of emissions})}{(\text{total time})(\text{sec})} = 0.8396$ 次/秒，紀錄時間的準確度可達 0.0002 秒。

在單位時間內所放射的 α 粒子數目為一隨機變數，故需先決定想要配適的分佈。可以假設：ⓘ 在觀察時段中，每個原子其發射速率為常數；

ⓘⓘ 觀察的α放射粒子數目，來自於很多個獨立的來源(原子)。

由上面的假設 ，Berkson 猜測α放射粒子的 粒子數 應具有卜瓦松分佈 (Poisson distribution)。因此，將所要做的假設檢定寫下來為：

H_0：數據來自於 Poisson distribution　versus　H_1：數據非為 Poisson distribution

然而，在開始做檢定之前，要先謹慎的確認所取得的資料是否有滿足 Poisson 分佈的基本假設，否則所做出來的檢定結果將不具任何意義。

Poisson 分佈的三個公設為：

① 在單位時間(或空間)內，事件發生的次數為一常數；

② 在沒有交集的時間(或空間)中，事件發生之間是相互獨立的；

③ 在同一時間點中，不會同時發生兩件事。

看起來所得到的數據皆滿足上面的假設。下表為觀察所得的資料與分析：

n	Observed	Expected	χ^2
0~2	18	12.2	2.76
3	28	27.0	0.04
4	56	56.5	0.01
5	105	94.9	1.07
6	126	132.7	0.34
7	146	159.1	1.08
8	164	166.9	0.05
9	161	155.6	0.19
10	123	130.6	0.44
11	101	99.7	0.02
12	74	69.7	0.27
13	53	45.0	1.42
14	23	27.0	0.59
15	15	15.1	0.00
16	9	7.9	0.57
17+	5	7.1	0.57
	1207	1207	8.99

上表以 10 秒為時間間隔,觀測每個間隔內放射的α粒子數,總共觀察了 1207 個時間間隔；第一欄為 10 秒內α放射粒子的個數，共區分成 16 種情況；第二欄數據代表的意思是，在每個情況中，所觀測到α放射粒子的個數；例如，第二列所代表的意思是，在 10 秒的間隔中α放射粒子的個數為 3，在 1207 個觀測間隔中，總共發生了 28 次。第三欄是假設若α放射粒子數為 Poisson 分佈時，所得的觀測個數。首先，先計算出每個情況所發生的機率：

$$\pi_k = P(x=k) = \frac{\lambda^k e^{-\lambda}}{k!}$$

其中λ是前面所述的平均放射速率 8.392 (次/10 秒)。需特別注意到的是第一列的機率為：

$$p_1 = \pi_0 + \pi_1 + \pi_2$$

而最後一列(第 16 列)的機率為：

$$p_{16} = \sum_{k=17}^{\infty} \pi_k$$

接著再將個別的機率值乘上總次數 1207，即可得到每種情況的期望個數。

最後一欄為卡方檢定的統計量,是由每種情況的觀測值個數（O_k）與期望值個數（E_k）計算而來的，方法如下：

$$\chi_k = \frac{(O_k - E_k)^2}{E_k}$$

因此，可以寫出卡方檢定的統計量為：

$$\chi^2 = \sum_k \frac{(O_k - E_k)^2}{E_k} = 8.99$$

統計量χ^2被證明出具有卡方分佈，自由度為：

自由度	=	所區分的情況個數	-	估計的獨立參數個數	- 1
	=	16	-	1	- 1
	=	14			

因此：

$$\chi^2 = \sum_k \frac{(O_k - E_k)^2}{E_k} \sim \chi^2_{14}$$

在此分佈下，統計量$\chi^2 = 8.99$所發生機率值 (p-value) 為：

$$\text{p-value} = P(\chi^2 > 8.99 \mid H_0 \text{ is true}) = 0.83$$

則在顯著水準(型 I 誤差)為 0.05 的假設下，因為 p-value = 0.83 > 0.05，因此我們沒有足夠證據拒絕虛無假設，即相信α放射粒子的個數是來自 Poisson 分佈。

參考文獻

Rice, J. A. (1995). *Mathematical Statistics and Data Analysis, second edition.* Wadsworth Publishing Company, Belmont, California.

Berkson, J. (1966). *Examination of randomness of alpha particle emissions.* In *Research Papers in Statistics*, F. N. David (ed.). New York: Wiley.