

普松分佈及負二項分佈在颱風 降水頻率上之應用

戴治台 徐 鎔 馬汝安

Application of the Poisson and Negative-Binomial Models to the Frequency of Typhoon and Precipitation

Chih-Tai Tai, Kae Shyu, Ruu-Ann Ma,

Abstract

In any quantitative analysis, one of the most important procedures is to find the probability distribution models of these random variables. Using Poisson distribution function and Negative-Binomial distribution function as models, we read in (i) the number of typhoon occurred in west Pacific Ocean area (110°E ~ 160°E , 5°N ~ 55°N) during the period from 1892 to 1971. (ii) the number of typhoon which invaded Taiwan during the same period. (iii) the number of precipitation days and thunderstorm days at Taipei (1897~1970) and Ilan (1936~1970). And use Chi-Square distribution with $n-1$ degree of freedom at the 0.05 level of significance to test the validity of these models.

We find at least one of these two distributions can be used as model for these choised random variables at 95% confidence.

一、簡 介

一個科學的模型 (Scientific model) 是一個現象的抽象和簡要的描述，科學的模型常基於數學上的分析。如果模型和現象間的相似性足夠顯著，則我們對於模型的瞭解常可提供對有關現象之結果的解釋或更進一步的研究。例如對於某一特殊現象（如雷雨）建立了一個數學的模型，則吾人可將由這個現象所得的資料（如雷雨發生的次數）以數學模型中的變數值來代表，如果這種對應的關係足夠正確，則吾人不必再做煩複的實驗以獲得此現象的新事實，而可直接由這些數學模型出發，研究更複雜的大自然現象。所以在任何定量分析中，最重要的步驟之一是找出這些隨機變數（Random Variables）的或然率分佈模型。基於此，本文之目的乃在蒐集 (i) 西太平洋 (110°E — 160°E , 5°N — 55°N) 颱風在 80 年內 (1892—1971) 發生的總次數，(ii) 侵臺的颱風次數 (1892—1971年

)，以及 (iii) 臺北 (1897—1970)，宜蘭 (1936—1970)，降水日及雷雨日的發生次數。逐月利用普松分佈 (Poisson distribution) 和負二項式分佈 (Negative-Binomial distribution) 的或然率分佈來建立這些現象的數學模型，並利用卡方分佈 $\chi^2_{0.05, n-1}$ (Chi-Square distribution with $n-1$ degree of freedom at the 0.05 level of significance) 來試驗這些模型的可靠性。

二、模型的計算及試驗方法

1. 普松分佈的或然率函數

$$(1) f(x) = \frac{e^{-\lambda} \lambda^x}{x!}, x=0,1,2,\dots$$

式中 x 是某一特殊現象（如雷雨）每月或每年發生的日數， $x=0$ 代表無此現象之發生， λ 是這些樣本的平均值 (Sample mean)。

2. 負二項分佈的或然率函數

$$(2) f(x;p,r) = \binom{x+r-1}{r-1} p^r (1-p)^x$$

$$= \frac{(x+r-1)!}{x! (r-1)!} p^r (1-p)^x$$

$x=0, 1, 2, \dots$

p, r 是這個或然分佈的參數。

x 代表的意義與前中所述者相同。

吾人可證明，利用矩法 (moment method) 估計出之參數值如下：

$$(3) \quad r = \frac{\mu_x^2}{\sigma_x^2 - \mu_x}, \quad p = \frac{\mu_x}{\sigma_x^2}$$

上式 μ_x 是樣本的平均值

σ_x^2 是樣本的變方 (Sample Variance)

3. 設在第 i 組 (class) 中真正觀察到的某一現象的數目有 y_i 次，且對一所有的 n 個觀察，由模型算出期望落入第 i 組的次數有 $n f_i$ 次，則

$$\chi^2 = \sum_{i=1}^k \frac{(y_i - n f_i)^2}{n f_i} \quad \text{可近似滿足卡方分佈，式}$$

中， k 是整理出的組數， $n f_i$ 至少為 5，本文中取 $n f_i=8$ ，此用來試驗：

(i) 零假設 (null hypothesis) H_0 ：現象的發生是滿足等均值 (equal mean) 的普松分佈模型。

佈模型。

(ii) 對立假設 H ：現象的發生不滿足等均值的普松分佈模型，而是滿足負二項分佈模型。

如果由零假設算出來的值 $\chi^2_{K-1} \geq \chi^2_{0.05, K-1}$ ，($\chi^2_{0.05, K-1}$ 表 $k-1$ 個自由度的資料在 5% 顯著水準下的臨界值)。則吾人摒棄此零假設而採用對立的假設，並用同樣的方法試驗其可靠性。

在程式裏我們亦加下面兩個手續以增進試驗的可靠性。

① 自動重組觀測的數據以保證每組 $n f_i \geq 8$ ，如果 $n f_i < 8$ 則下一組或下幾組將合併起來以保證每一組 $n f_i \geq 8$ ，(但若某一現象每年或每月只發生一次或全不發生者不再做此合併)。

② 利用葉氏修正法 (Yate's Correction)，對觀測值少 (多) 於模型期望值者加上 (減去) $\frac{1}{2}$ 。

$$(4) \quad \chi^2 = \sum_{i=1}^K \frac{(|y_i - n f_i| - \frac{1}{2})^2}{n f_i}$$

k 是利用①歸併後的組數，以彌補不連續的分佈用連續的卡方分佈近似，所生的誤差。

4. 主要計算過程如圖 1 方塊圖所示

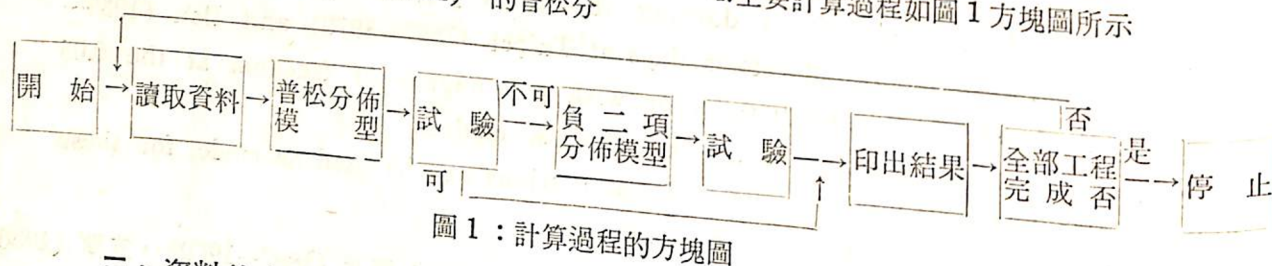


圖 1：計算過程的方塊圖

三、資料的來源及計算結果

(i) 我們讀取「中央氣象局」編印的「八十年颱風路徑圖」， $110^\circ\text{E} \sim 160^\circ\text{E}$ ， $5^\circ\text{N} \sim 55^\circ\text{N}$ 西太平洋區內颱風的次數及侵臺次數，逐月利用上節的說明，所得之結果如表(1)及表(2)。

表(1)：西太平洋區 ($110^\circ\text{E} \sim 160^\circ\text{E}$ ， $5^\circ\text{N} \sim 55^\circ\text{N}$) 颱風生成次數的逐月模型 (1892~1971年)

月	風或然率模型	自由度	χ^2_{K-1}	$\chi^2_{0.05, K-1}$
1	P	(1)	0.180×10^0	(3.8410)
2	P	(1)	0.165×10^1	(3.8410)
3	P	1	0.646×10^{-3}	3.8410
4	P	1	0.427×10^0	3.8410
5	NB	1	0.924×10^0	3.8410

6	P	2	0.425×10^1	5.9910
7	P	4	0.109×10^1	9.4880
8	P	5	0.666×10^1	11.0700
9	P	5	0.114×10^1	11.0700
10	P	4	0.264×10^1	9.4880
11	NB	2	0.731×10^1	5.9910
12	P	2	0.284×10^1	5.9910
總合	NB	6	0.324×10^1	12.5910

表 1 中：P 表普松分佈模型

NB 表負二項分佈模型

χ^2_{K-1} 表在所用的模型假設下的卡方值

$\chi^2_{0.05, K-1}$ 見上節中之說明

以下各表之符號意義均同此處。

表(2) 颱風侵臺次數的逐月模型 (1892~1971年)

月	或然率模型	自由度	χ^2_{K-1}	$\chi^2_{0.05, K-1}$
1	P	(1)	0.244×10^0	(3.8410)
2	P	(1)	0.551×10^1	(3.8410)
3	P	(1)	0.244×10^0	(3.8410)
4	P	1	0.367×10^1	3.8410
5	P	1	0.701×10^{-2}	3.8410
6	P	1	0.646×10^{-3}	3.8410
7	P	2	0.900×10^{-1}	5.9910
8	P	2	0.249×10^1	5.9910
9	P	2	0.306×10^0	5.9910
10	P	1	0.379×10^{-1}	3.8410
11	P	1	0.177×10^{-2}	3.8410
12	P	1	0.244×10^0	3.8410
總合	P	5	0.208×10^1	11.0700

(ii) 我們又從氣象局編印之臺灣歷年觀測報告中讀取臺北 (1897~1970)，宜蘭 (1936~1970) 雷雨、降水日數。其結果如表(3)，表(4)，表(5)，表(6)。

表(3) 臺北雷雨日數的逐月模型 (1897~1970)

月	或然率模型	自由度	χ^2_{K-1}	$\chi^2_{0.05, K-1}$
1	P	1	0.108×10^1	3.8410
2	NB	1	0.758×10^0	3.8410
3	NB	3	0.308×10^1	7.8150
4	NB	3	0.228×10^1	7.8150
5	P	5	0.509×10^1	11.0700
6	P	5	0.851×10^1	11.0700
7	NB	4	0.196×10^1	9.4880
8	P	6	0.584×10^1	12.5920
9	NB	4	0.416×10^1	9.4880
10	P	2	0.264×10^1	5.9910

11	P	1	0.705×10^0	3.8410
12	P	1	0.121×10^{-2}	3.8410

表(4) 臺北降水日數的逐月模型 (1897~1970)

月	或然率模型	自由度	χ^2_{K-1}	$\chi^2_{0.05, K-1}$
1	P	5	0.777×10^1	11.0700
2	P	5	0.452×10^1	11.0700
3	P	5	0.131×10^1	11.0700
4	NB	4	0.400×10^1	9.4880
5	P	5	0.155×10^1	11.0700
6	P	5	0.396×10^1	11.0700
7	P	5	0.129×10^1	11.0700
8	P	5	0.134×10^1	11.0700
9	P	5	0.538×10^1	11.0700
10	P	5	0.179×10^1	11.0700
11	P	5	0.116×10^1	11.0700
12	NB	4	0.588×10^1	9.4880

表(5) 宜蘭雷雨日數的逐月模型 (1936~1970年)

月	或然率模型	自由度	χ^2_{K-1}	$\chi^2_{0.05, K-1}$
1	P	1	0.365×10^1	3.8410
2	P	1	0.240×10^{-2}	3.8410
3	P	2	0.168×10^1	5.9910
4	P	2	0.543×10^0	5.9910
5	P	2	0.267×10^1	5.9910
6	P	2	0.153×10^1	5.9910
7	NB	1	0.314×10^{-1}	3.8410
8	P	2	0.642×10^0	5.9910
9	P	2	0.203×10^1	5.9910
10	P	1	0.205×10^{-1}	3.8410
11	P	1	0.381×10^{-2}	3.8410
12	P	1	0.236×10^0	3.8410

表(6) 宜蘭降水日的逐月模型 (1936~1970年)

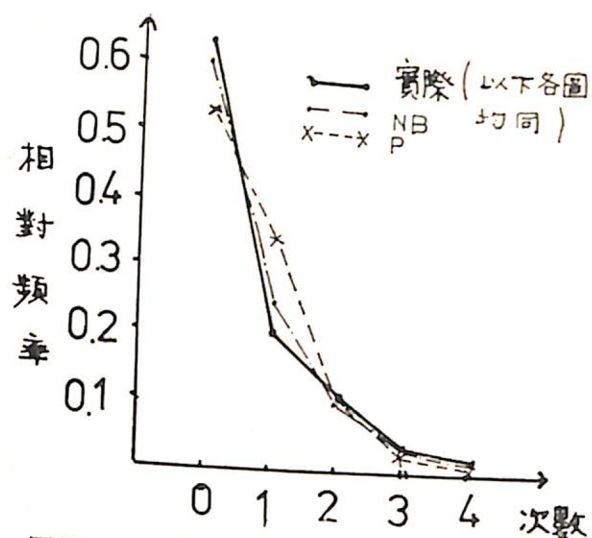
月	降水	或然率模型	自由度	χ^2_{K-1}	$\chi^2_{0.05, K-1}$
1	P	P	2	0.214×10^1	5.9910
2	P	P	2	0.463×10^0	5.9910
3	P	P	2	0.372×10^0	5.9910
4	P	P	2	0.211×10^1	5.9910
5	P	P	2	0.517×10^1	5.9910
6	P	P	2	0.148×10^1	5.9910
7	P	P	2	0.953×10^{-2}	5.9910
8	P	P	2	0.415×10^{-1}	5.9910
9	P	P	2	0.144×10^0	5.9910
10	P	P	2	0.172×10^1	5.9910
11	P	P	2	0.131×10^1	5.9910
12	P	P	2	0.475×10^1	5.9910

四、結 論

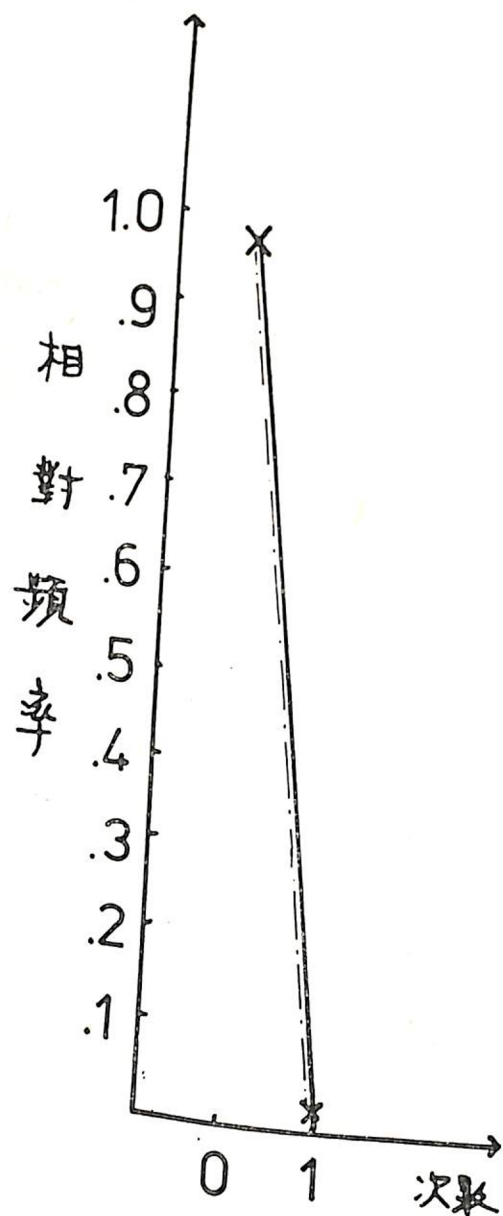
1. 本文採用5%的顯著水準 (level of significance), 所以得出各模型正確性有 95%。
2. 普松分佈模型的計算比較簡單, 故我們先試驗其可靠性。若模型可靠性無法達到95%, 則改用負二項分佈模型計算。表(7), 圖(2)便是此種情況的一個例子。

表(7) 西太平洋區 (110°E~160°E 5°N~55°N) 颱風生成次數 (1892~1971年, 各年五月的資料)。

實際觀察的次數	頻率	模型P算出的頻率	模型NB算出的頻率
0	50	0.423×10^2	0.482×10^2
1	16	0.270×10^2	0.197×10^2
2	9	0.859×10^1	0.757×10^1
3	3	0.183×10^1	0.284×10^1
4次及4次以上	2	0.291×10^0	0.105×10^1



圖(2) 西太平洋區颱風生成相對頻率 (1892~1971每年五月的資料) 與普松, 負二項式或然率分佈模型之比較。



圖(3) 颱風侵臺相對頻率與普松或然率分佈模型之比較。 (1892~1971年二月份)

表(9) 西太平洋區 (110°E~160°E 5°N~55°N) 颱風生成次數 (1892~1971年九月份)。

實際觀察次數	頻率	模型P算出的頻率
0	1	0.257×10^1
1	11	0.884×10^1
2	14	0.152×10^2
3	15	0.174×10^2
4	19	0.150×10^2
5	10	0.103×10^2
6	5	0.589×10^1
7次及7次以上	5	0.289×10^1

蘭的雷雨日 (除 7 月份外) 全年均可採普松分佈模型。而臺北市二、三、四月份雷雨日必需採負二項分佈模型。

5. 七、八、九月份在西太平洋發生和侵臺之颱風次數的或然率, 列於表(10)。各月份模型值與實際值的比較示於圖(5), 圖(6), 圖(7)。

3. 普松分佈模型本身常是一個既簡單又適用的模型, 對於颱風的頻率可以得到良好之結果。表(8) 圖(3)和表(9)圖(4)是兩個例子。

表(8) 二月份颱風侵臺次數的模型 (1892~1971年)

實際觀察次數	頻率	模型P算出的頻率
0	79	0.790×10^2
1次及1次以上	1	0.988×10^0

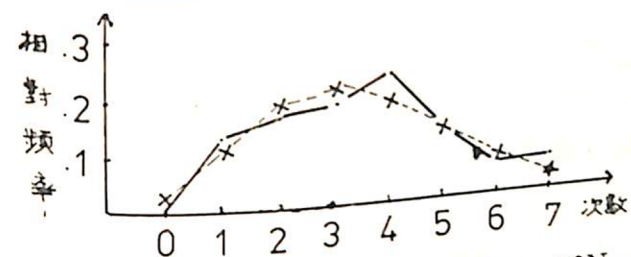


圖 4 : 西太平洋區 (110°E~160°E, 5°N~55°N) 九月份颱風生成相對頻率與普松或然率分佈模型之比較。

4. 颱風侵襲臺灣的次數, 以及宜蘭的降水, 宜

表(10) 西太平洋發生颱風次數及颱風侵臺次數的或然率 (1892~1971年)

次 數	七 月 份		八 月 份		九 月 份	
	西 太 平 洋	侵 臺	西 太 平 洋	侵 臺	西 太 平 洋	侵 臺
0	0.0715	0.478	0.0179	0.359	0.0321	0.472
1	0.189	0.353	0.0719	0.368	0.110	0.354
2	0.249	0.130	0.145	0.188	0.190	0.133
3	0.219	0.032	0.194	0.0644	0.218	0.0332
4	0.144		0.157	0.0165	0.137	
5	0.0765		0.105		0.129	
6	0.0334		0.0607		0.0737	
7	0.0126		0.050		0.0362	
8	0.00415					

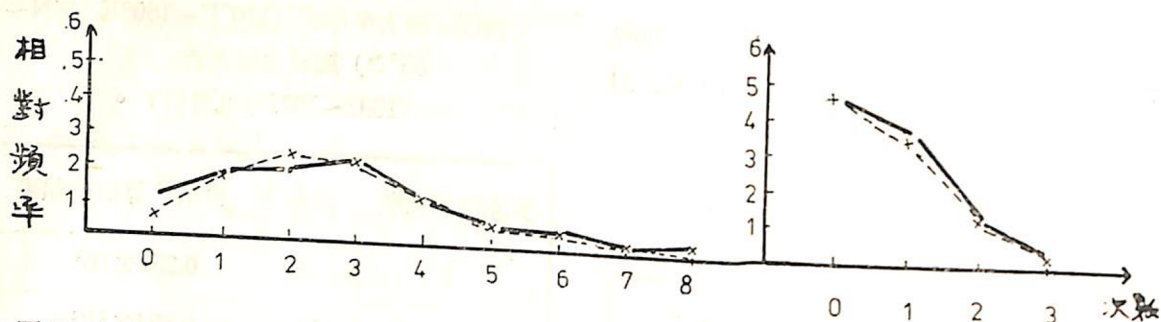


圖 5：(a)七月份西太平洋颱風生成相對頻率，(b)侵臺，與普松或然率分佈模型的比較。

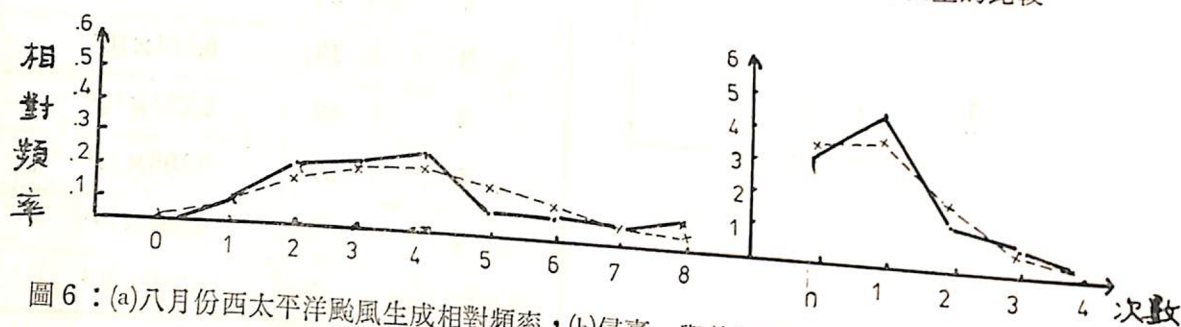


圖 6：(a)八月份西太平洋颱風生成相對頻率，(b)侵臺，與普松或然率分佈模型的比較。

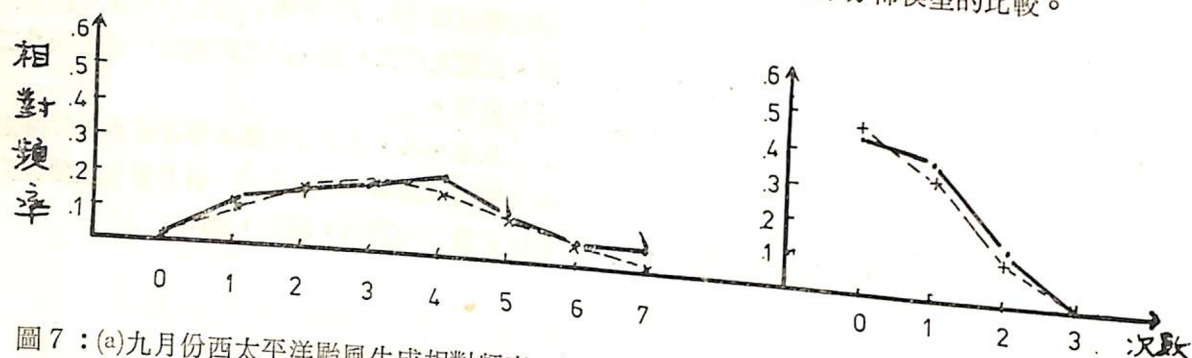


圖 7：(a)九月份西太平洋颱風生成相對頻率，(b)侵臺，與普松或然率分佈模型的比較。

誌謝：本文承空軍氣象中心俞家忠主任垂詢和關懷，長期課王時鼎課長的建議和鼓勵，以及張儀峰，林清榮等諸同仁協助讀取資料，在此一併致謝。

參考文獻

- (1) Carnahan: Applied Numerical

Methods.

- (2) F. James Rohlf and Robert R. Sokal: Statistical Table.
- (3) R. G. Bany and A. H. Perry: Synoptic Climatology, method and applications