銘 傳 大 學

應用統計與資料科學學系時間數列分析期末報告

101 觀景台來客量

學生:05170021 黃育榆

05170242 彭珺瑋

指導老師:廖承茂 博士

目錄

圖目錄	2
第一章研究動機與目的	3
第二章資料來源與解釋	3
第三章研究方法與步驟	5
第四章 研究過程與結果分析	5
第五章 研究結論與討論	11
第六章 參考資料	11

圖目錄

啚	2-1:原始資料表格	3
昌	2-2:整理後表格	4
昌	2-3: 要删除的表格	4
邑	3-1:研究流程圖	5
昌	4-1:時間數列圖與其ACF、PACF	6
置	4-2:單位根檢定圖	6
昌	4-3: 差分後的時間數列圖與其 ACF、PACF	7
昌	4-4:差分後的單位根檢定圖	7
昌	4-5:ARIMA(3,1,0)(2,1,0) ₁₂ 圖	8
昌	4-6:ARIMA(3,1,0)(3,1,0) ₁₂ 圖	8
昌	4-7:ARIMA(3,1,0)(4,1,0) ₁₂ 圖	9
昌	4-8:AIC(3,1,0)(2,1,0)圖	9
昌	4-9:AIC(3,1,0)(3,1,0)圖	9
昌	4-10:AIC(3,1,0)(4,1,0)圖	9
昌	4-11: 最佳模型係數圖	10
圖	4-12:模型預測圖	10

第一章 研究動機與目的

一說到台灣最有名的景點之一就一定會有臺北的 101 大樓,每年都會有許多觀光客前去參觀遊玩,而其中能登上大樓頂端觀景台眺望整個臺北市,是許多觀光客來到 101 大樓的主要目的之一,頂樓觀景台是需要先買票才能進去的地方,且擁有絕佳 360°全視野景觀並提供多項服務,其中還可以看到世界最大,最重,唯一外露供觀賞的風阻尼球。

根據臺北旅遊網的主要觀光遊憩區遊客人數調查,可以得知從 2005 年 1 月到 2020 年 1 月到臺北 101 大樓的觀景台總人數達 2931 萬人。因此我們希望透過此次報告,用臺北 101 觀景台 94 年 1 月至 109 年 1 月的來客數,推估接下來來客量的走勢。

第二章 資料來源與解釋

- 1. 臺北 101 觀景台:臺北 101 大樓的 89 樓室內觀景台、91 樓爲室外觀景臺。
- 資料範圍:共 181 筆(94 年 1 月至 109 年 1 月),一個月一筆,共 15 年又 1 個月
- 3. 資料來源:臺北旅遊網(https://www.travel.taipei/zh-tw)
- 4. 原始資料表格:

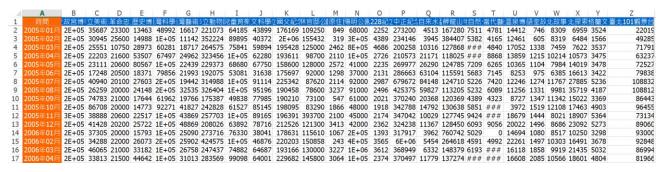


圖 2-1 原始資料表格

5. 整理後表格:

	Α	В	С
1	time	臺北101體	見景台
2	2005年01	22019	
3	2005年02	49285	
4	2005年03	71791	
5	2005年04	63237	
6	2005年05	72527	
7	2005年06	79838	
8	2005年07	108832	
9	2005年08	108812	
10	2005年09	86443	
11	2005年10	96455	
12	2005年11	73134	

圖 2-2 整理後表格

變數內容說明

(1) time: 時間(xx 年 xx 月)

(2) 臺北 101 觀景台:當月來觀景台的總人數

P.S.整理後的表格是把其他地區的來客人數刪掉

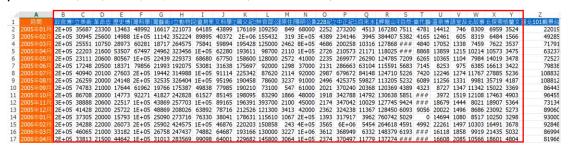


圖 2-3 要刪除的表格

第三章 研究方法與步驟



第四章 研究過程與結果分析

- 1. 安裝及載入時間數列套件
- 2. 將資料讀入 R
- 3. 畫出時間數列圖與其 ACF、PACF

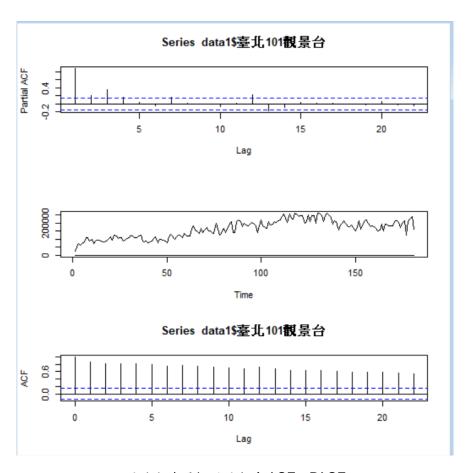


圖 4-1 時間數列圖與其 ACF、PACF

Augmented Dickey-Fuller Test

data: datal\$臺北101觀景台

Dickey-Fuller = -1.9236, Lag order = 5, p-value = 0.6082

alternative hypothesis: stationary

圖 4-2 單位根檢定圖

H_0:數列含有單位根(數列為非平穩型)

H_1:數列為平穩型

 $\alpha = 0.05$

p-value=0.6082

::0.6082 > 0.05

∴.不拒絕 H_0,表數列有單位根存在,即為非平穩型時間數列,因此進一步做一階差分

4. 進行一次差分,並畫出差分後的時間數列圖與其 ACF、PACF

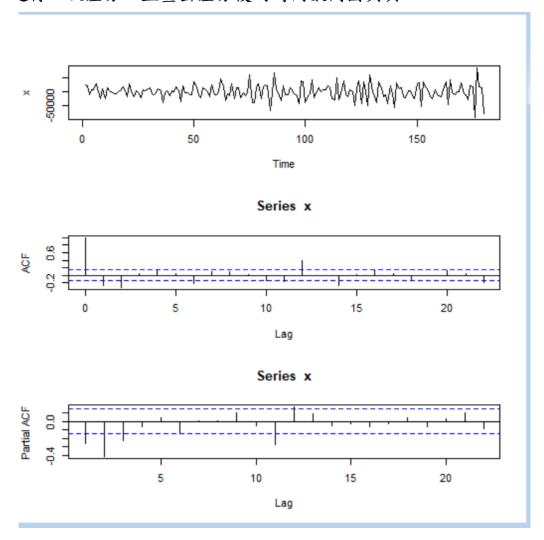


圖 4-3 差分後的時間數列圖與其 ACF、PACF

由圖 4-3 的第一張時間數列圖可以看出此時間數列上下波動越來越大趨勢。

由圖 4-3 的第二張 ACF 圖可以看出在 lag=1,2,12 時較高其餘大多落在範圍內,此時間數列有呈現季節性的變化。

由圖 4-3 的第三張 PACF 圖可以看出此時間數列在 lag=1,2,3,11,12 時較高其餘大多落在範圍內。

Augmented Dickey-Fuller Test

data: x
Dickey-Fuller = -7.9347, Lag order = 5, p-value = 0.01
alternative hypothesis: stationary

Warning message:
In adf.test(x): p-value smaller than printed p-value

圖 4-4 差分後的單位根檢定圖

H_0:數列含有單位根(數列為非平穩型)

H_1:數列為平穩型

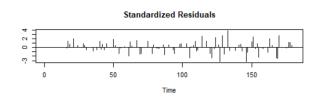
 $\alpha = 0.05$

p-value=0.01

∵0.01 < 0.05

∴拒絕 H_0,表數列無單位根存在,即為平穩型時間數列

5. 季節性模型



ACF of Residuals



p values for Ljung-Box statistic

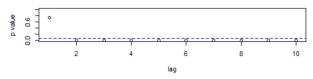


圖 4-5 ARIMA(3,1,0)(2,1,0)₁₂ 圖

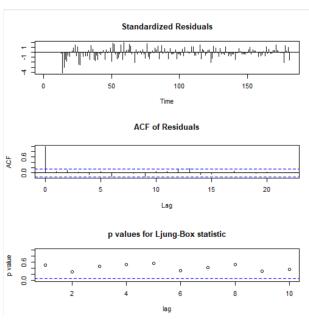


圖 4-6 ARIMA(3,1,0)(3,1,0)₁₂ 圖

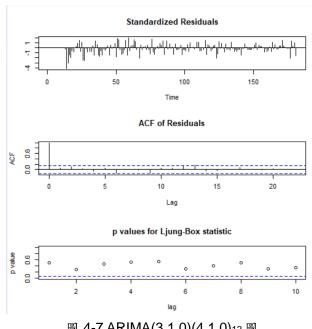


圖 4-7 ARIMA(3,1,0)(4,1,0)₁₂ 圖

由圖 4-5,4-6,4-7 中 ACF 圖都有落在兩條藍色線內代表適當的圖

6. 模型選取(比較 AIC)

圖 4-10 AIC(3,1,0)(4,1,0)圖

由圖 4-8,4-9,4-10 可以看出圖 4-9 AIC(3,1,0)(3,1,0)的 AIC 為最小,因此選它作為我們的最佳模型圖。

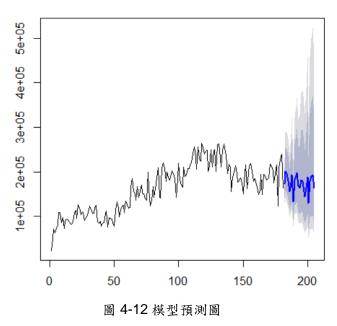
7. 最佳模型

刪除不符合的模型後,發現以季節性差分 MA(3)的 AIC 為最小,因此決定最佳模型為(3,1,0)(3,1,0),其係數值如下:

圖 4-11 最佳模型係數圖

8. 模型預測

Forecasts from ARIMA(3,1,0)(3,1,0)[12]



如圖 4-12 中,黑色曲線部分為 2005 年 1 月至 2020 年 1 月的 181 筆原始資料,藍色實線為根據先前資料所做出的預測未來 3 年資料走向,較深的藍色為 95%區間,淺藍色為 80%區間,由此可以看出來未來 3 年的預測資料皆位於 95%區間內,表示符合預期。

第五章 研究結論與討論

我們透過每 12 個月一個循環,這種季節性的方式,找到最佳平穩型的模型。 而最後使用 ARIMA(3,1,0)(3,1,0)12 為我們的最佳模型,並利用此模型對我們的 181 筆原始資料做預測,根據研究結果顯示,在 2005 年 1 月至 2020 年 1 月中, 可發現每年的 4、7、8、12 月人數顯示為較多,推知可能是 4 月的清明連假、7、 8 月的暑假,以及 12 月的年底休假與聖誕節的關係使得觀光人數明顯增多,而 每年的 6、9 月的人數相對的減少了許多,我們推論可能是因為 6 月的梅雨季及 9 月的開學季的關係使得觀光人數明顯減少。由此研究結果我們認為在 2、3、9 月的時候到 101 大樓觀景台參觀不會有出現人擠人的情況,同時也能好好欣賞 風景。

第六章 參考資料

- 臺北旅遊網
 https://www.travel.taipei/zh-tw
- 2. 首頁| Taipei 101 https://www.taipei-101.com.tw/tw