

時間序列分析 期末報告

107508006 歐語四 陳葳芃 108405056 廣告四 楊岳鋁

報告大綱

一、研究目的

二、資料介紹

Time series、ACF、PACF

三、原始資料探索

時間序列觀察 Raw data ACF PACF 一階差分後的比較

四、不使用差分來 de-trend 進行 order 選取

參數估計 Estimation - 最高次係數是否異於 0? 是否接近 1?

模型診斷 Model diagnostic - 3 方法

模型修正 Model refinement?

五、使用差分來 de-trend 進行 order 選取

參數估計 Estimation - 最高次係數是否異於 0? 是否接近 1?

模型診斷 Model diagnostic - 3 方法

模型修正 Model refinement?

六、使用差分來 de-trend 進行 order 選取(含季節性差分)

參數估計 Estimation - 最高次係數是否異於 0? 是否接近 1?

模型診斷 Model diagnostic - 3 方法

模型修正 Model refinement?

七、模型選擇

八、預測

九、結論

十、附錄——參考文獻及資料來源

一、研究目的

台灣逐漸邁入高齡化社會，也鑒於此分析台灣出生人口時間序列的趨勢，探究出生人口的變遷與影響因素對於國家人口政策制定事關重要。

因此，我們想要瞭解每月出生人口的時間變動的關聯性或週期性，同時，深入探討背後的社會、文化和經濟背景，以促進找出有意義的資訊來給予我們的政府相關施政方針的建議。

二、資料介紹

本次報告研究的資料集為台灣月出生人口統計，來源自中華民國統計資訊網——總體統計資料庫，其中紀錄自 2000 年 1 月以來每月出生人數，資料目前更新至 2023 年 3 月。樣本個數為 279 筆，變數個數為 3 個。

Table.1 資料型態及前六筆資料

	time	amount	month
類別	時間 time	數值 numerical	類別 categorical
說明	年月份	出生人數	月份
	2000-01-01	24042	1
	2000-02-01	22263	2
	2000-03-01	26259	3
	2000-04-01	22587	4
	2000-05-01	25019	5
	2000-06-01	22937	6

三、原始資料探索

首先，我們先將資料原始 Time series 給畫出來如 Figure.1，可以看出整體生育人數是在往下降的，接著我們好奇哪些月份生育率會特別高？而哪些月份又特別低？又或者是哪些異於常理的值出現在我們的資料及當中。

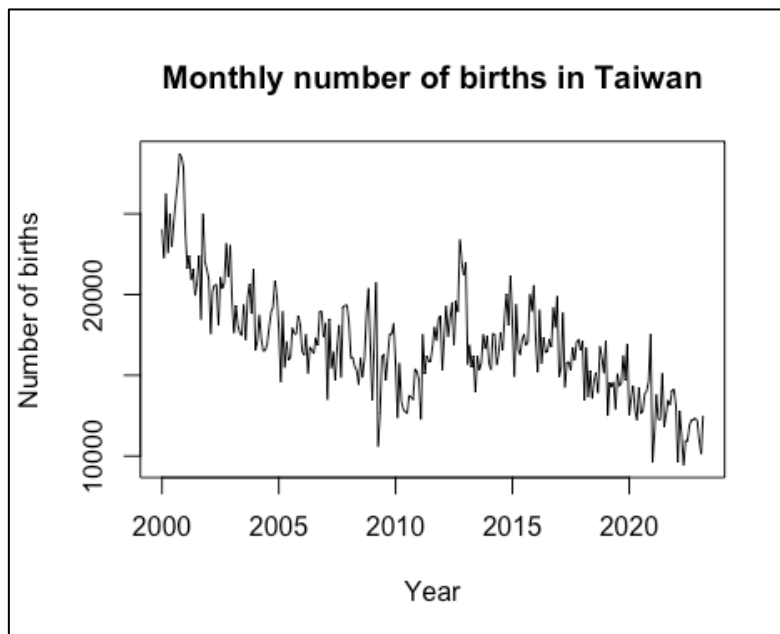


Figure.1 原始出生人數時間序列圖

我們使用 `ggplot` 將每個點標上月份，如 Figure.2 所示，此時可以發現每年的相對極大值幾乎都是發生在 10、12 月，故將生育率高峰的此 2 個月特別用顏色標注起來。除此之外，我們也可以發現 2 個異常之處：

- (1) 2009 年 3 月為一個分水嶺，4 月開始急遽下降且 2010 年整年生育率呈現一個低迷的狀況
- (2) 2012 年 生育率出現拉升的情況，違反剛開始觀測的長期觀測趨勢

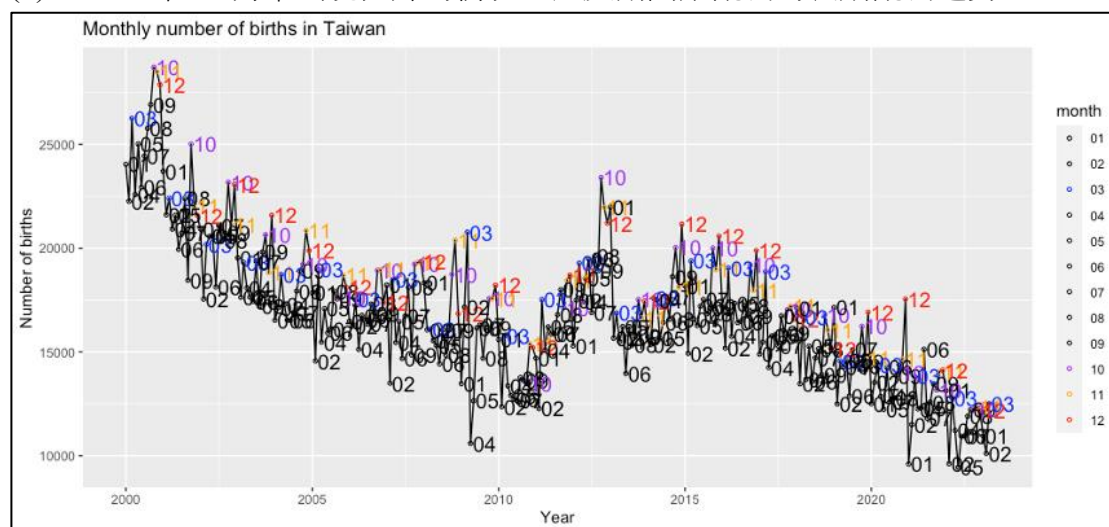


Figure.2 使用 `ggplot` 將時間序列圖標注月份

細究第 1 點原因，我們想到有可能是早些年發生的全球金融海嘯風暴，並透過論文來驗證假設。根據戴庭玉於 2011 在陽明交通大學的研究論文《美國 2007 年至 2009 年引發環球金融風暴的制度因素》第七頁指出貨幣流通速度與 2008 年第三季來到最高，象徵市場蕭條來臨。另一方面，王英傑於 2015 年中央大學完成的研究論文《景氣循環對婚育行為之影響》第八頁提及：『如過人們認為景氣蕭條持續，那對將來預期財富是下降的，.....，造成生育率下降。』

加上常理來說預產期懷胎 10 月左右的先備知識來看，若 2008 年第三季停止懷孕剛好會對應上 2009 年 3 月，也就是說以上兩篇論文都可以佐證該出生人口下降分水嶺的很好原因。

關於第 2 點，2012 年生育率不但恢復，而且更甚經濟蕭條前的數字。根據時任行政院政務委員薛承泰的採訪表示，政府於該段時期前的虎年就開始鼓勵結婚生育，並鼓吹年輕夫婦民國 99 年「愛情久久」年訂定婚約的風潮、生個民國百年寶寶或再下個年度生個龍寶寶，表示應是生育率竄高原因之一。

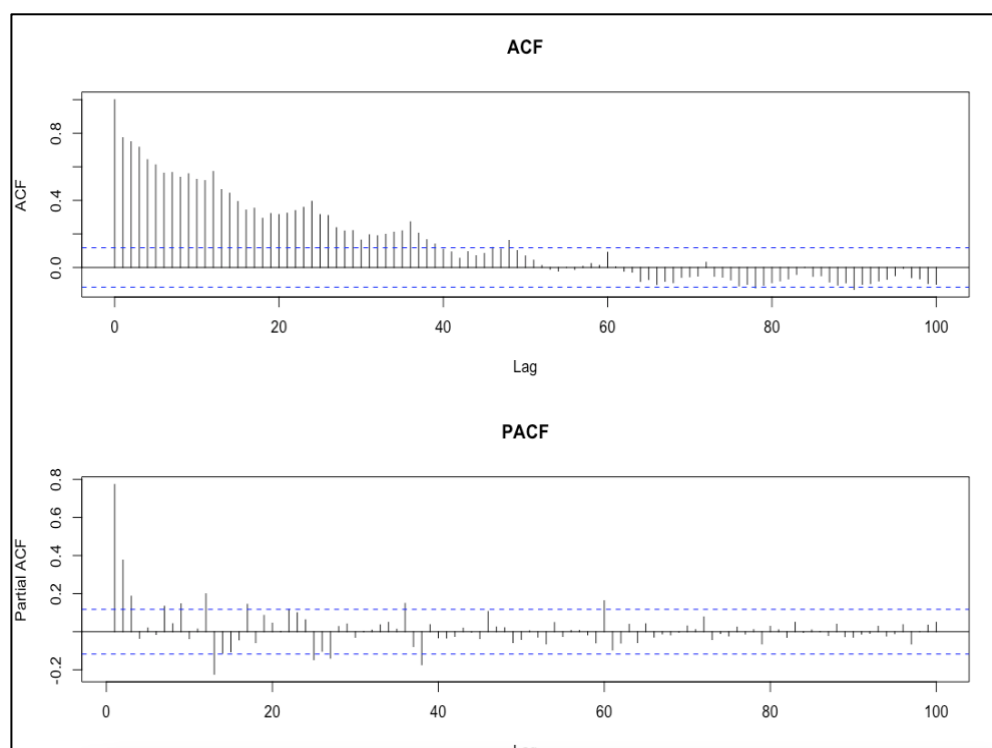


Figure.3 原始出生人數 ACF、PACF 圖

探索資料的時間序列圖結束後，藉由 Figure.3 原始資料 ACF PACF 判斷出我們此筆資料，ACF 值下降的不算緩慢，但同時也沒有明顯的快速下降，關於去除趨勢的一項抉擇——要不要差分，則是在經驗法則判斷的邊緣。因此我們決定兩種方式都進行，最後再從中選取比較好的模型進行預測。

除了藉由 Figure.3 判斷之外，我們亦透過 ADF 檢定，來檢驗此資料是不是屬於 stationary。此檢定虛無假設為此序列不是 stationary，對立假設為此序列為 stationary。以下為檢定原始資料的結果：

Augmented Dickey-Fuller Test

```
data: data
```

```
Dickey-Fuller = -3.6269, Lag order = 6, p-value = 0.03096
```

```
alternative hypothesis: stationary
```

根據檢定結果，給定 $\alpha=0.05$ 的情形下，我們可以拒絕虛無假設並宣稱此筆資料屬於 stationary。

四、不進行差分直接來進行 order 選取

參考 Figure.3 的結果，及前面所做的 ADF test，判斷 stationary 狀況。

【Order 選取與參數估計】

判斷：具有週期為 12 的季節性、ACF 視 tails off 或 cuts off at lag=4 or 5、

PACF cuts off at lag=2, 3 or 5

ACF 前後有約遞減的趨勢，非季節性部分則考慮 AR 最多取到 4 or 5

Table.2 模型選取與其相對應參數係數(不差分)

	Ar1	Ma1	Sar1	Sar2	Sar3	Sar4	Sar5
SARIMA(0,0,0)x(4,0,0) ₁₂			0.7413	0.1331	0.1874	-0.1669	
s.e.			0.0654	0.0820	0.0804	0.0687	
SARIMA(0,0,0)x(5,0,0) ₁₂			0.7303	0.1535	0.1972	-0.1201	-0.0741
s.e.			0.0657	0.0837	0.0805	0.0812	0.0693
	Sar1	Sar2	Sar3	Sar4	Sar5	Sma1	Sma2
SARIMA(0,0,0)x(4,0,1) ₁₂	1.2983	-0.2607	0.132	-0.2238		-0.5796	
s.e.	0.1573	0.1520	0.112	0.0677		0.1559	
SARIMA(0,0,0)x(4,0,2) ₁₂	2.2193	-1.6740	0.3976	-0.0165		-1.6902	1
s.e.	0.0793	0.0159	0.1723	0.0823		0.0700	0.0708
SARIMA(0,0,0)x(3,0,1) ₁₂	0.0330	0.5912	0.2255			0.6973	
s.e.	0.1873	0.1586	0.0660			0.1857	

Table.3 模型選取與其相對應參數係數(不差分)

	AR 最高	MA 最高	Sar1	Sar2	Sar3	Sar4	Sma1
SARIMA (1,0,2)x(4,0,1) ₁₂	0.9887	0.2533	0.9171	-0.0847	0.1390	0.0230	-0.8702
s.e.	0.0115	0.0709	0.0810	0.0863	0.0856	0.0735	0.0663
SARIMA (1,0,1)x(4,0,1) ₁₂	0.9937	-0.6656	0.9423	-0.1157	0.1281	0.0394	-0.8818
s.e.	0.0067	0.0425	0.0778	0.0863	0.0871	0.0727	0.0636
SARIMA (0,0,2)x(3,0,1) ₁₂		0.3051	0.1445	0.3079	0.2906		0.4259
s.e.		0.0579	0.2265	0.1559	0.0687		0.2380
SARIMA (1,0,2)x(3,0,1) ₁₂	0.9885	0.2543	0.9281	-0.0878	0.1546		-0.8783
s.e.	0.0112	0.0706	0.0737	0.0862	0.0683		0.0605
SARIMA (1,0,1)x(3,0,1) ₁₂	0.9940	-0.6678	0.9586	-0.1217	0.1580		-0.8935
s.e.	0.0071	0.0425	0.0709	0.0863	0.0672		0.0584

【模型診斷- SARIMA(0,0,2)x(3,0,1)₁₂】

method.1 檢驗殘差是否符合常態

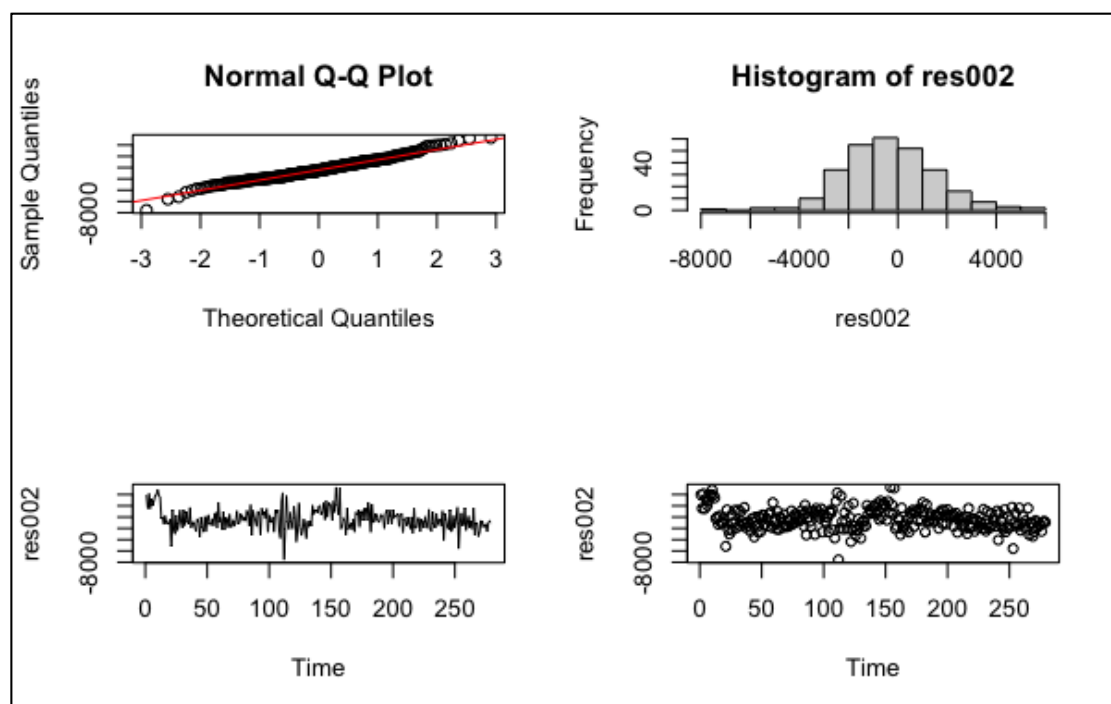


Figure.4 SARIMA(0,0,2)x(3,0,1)₁₂ 殘差常態檢驗

由 Figure.4 殘差檢驗，各方面判定其殘差應是在 0 附近跳動、但直方圖看起來也有左右對稱、由 Q-Q plot 判定僅接近符合常態。但為求謹慎，我們決定再進行 Shapiro-Wilk normality test，此檢定虛無假設為資料來自於常態，而對立假設則為資料不來自於常態，檢定結果如下：

Shapiro-Wilk normality test

data: res002

W = 0.98592, p-value = 0.007788

根據檢定結果，在 $\alpha=0.05$ 下，我們會 reject H_0 ，拒絕虛無假設：資料來自於常態，宣稱殘差是不符合常態分配的。

method.2 檢驗殘差是否相關

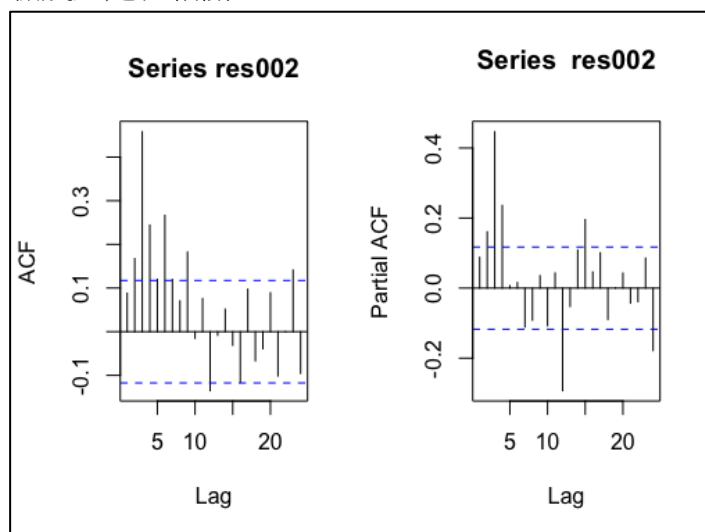


Figure.5 SARIMA(0,0,2)x(3,0,1)₁₂ 殘差 ACF 與 PACF 圖

根據 Figure.5 殘差的 ACF 與 PACF 圖，殘差大抵都不是在顯著線以下，顯示殘差間大致上可能尚存在相關性，所以假設誤差是 white noise 應是不可被接受的。

method.3 檢驗模型配飾是否適合

接著使用 Ljung Box test 來檢驗模型是否配飾得宜，此檢定

H_0 : the fitted model is adequate H_1 : the fitted model is not adequate

由 Figure.6 所示 Ljung Box test 均無通過，此模型並不是一個配適良好的選擇。

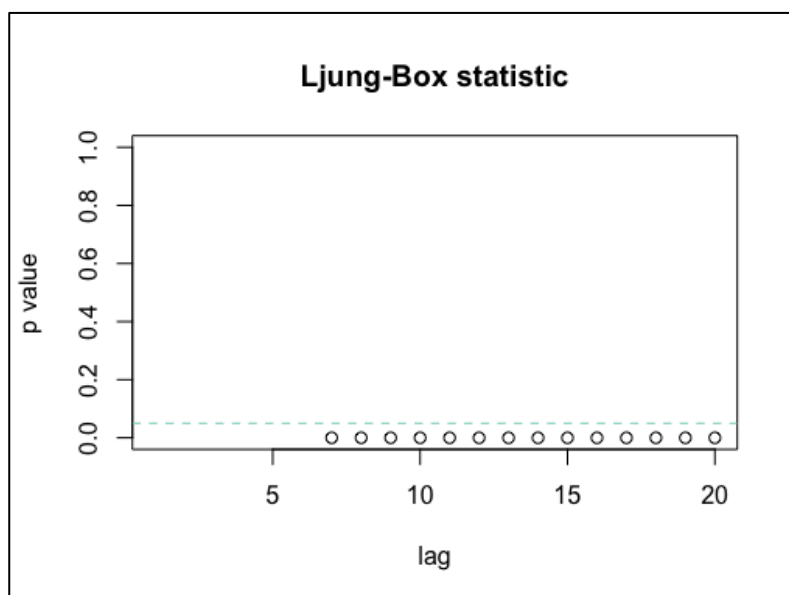


Figure.6 Ljung Box test for SARIMA(0,0,2)x(3,0,1)₁₂

【模型診斷- SARIMA(1,0,2)x(3,0,1)₁₂】

method.1 檢驗殘差是否符合常態

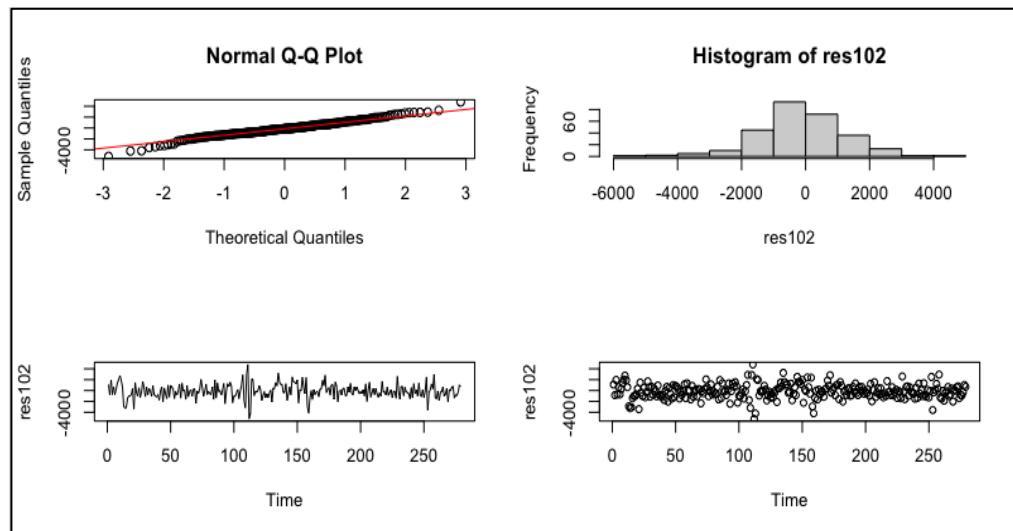


Figure.7 SARIMA(0,0,2)x(3,0,1)₁₂ 殘差常態檢驗

由 Figure.7 殘差檢驗，各方面判定其殘差沒有在 0 附近跳動、直方圖看起來也有左偏趨勢、由 Q-Q plot 判定不符合常態。但為求謹慎，我們決定再進行 Shapiro-Wilk normality test，此檢定虛無假設為資料來自於常態，而對立假設則為資料不來自於常態，檢定結果如下：

Shapiro-Wilk normality test

data: res102

W = 0.98509, p-value = 0.005344

根據檢定結果，在 $\alpha=0.05$ 下，我們會 reject H_0 ，拒絕虛無假設：資料來自於常態，宣稱殘差是不符合常態分配的。

method.2 檢驗殘差是否相關

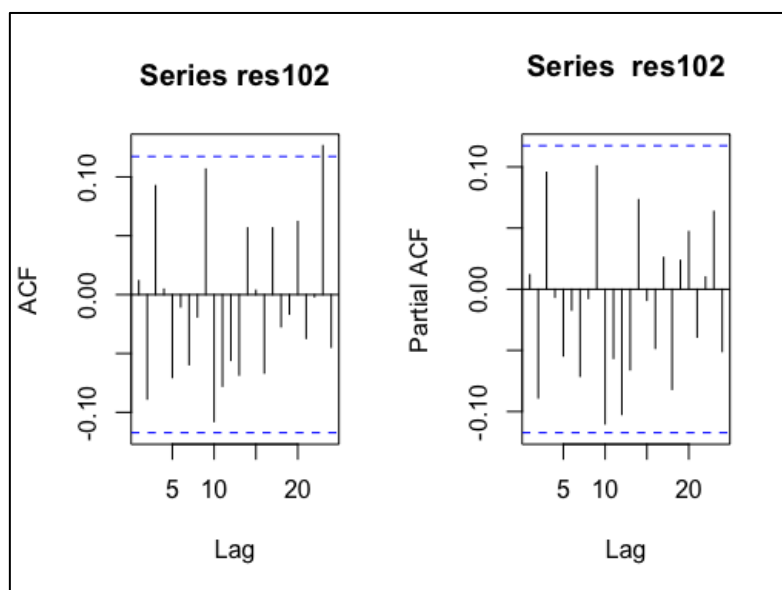


Figure.8 SARIMA(1,0,2)x(3,0,1)₁₂ 殘差 ACF 與 PACF 圖

根據 Figure.8 殘差的 ACF 與 PACF 圖，殘差幾乎是在顯著線以下，顯示殘差間大致上無相關性，所以假設誤差是 white noise 是可以被接受的。

method.3 檢驗模型配飾是否適合

接著使用 Ljung Box test 來檢驗模型是否配飾得宜，此檢定

H_0 : the fitted model is adequate H_1 : the fitted model is not adequate

由 Figure.9 所示 Ljung Box test 均無通過，此模型並不是一個配適良好的選擇。

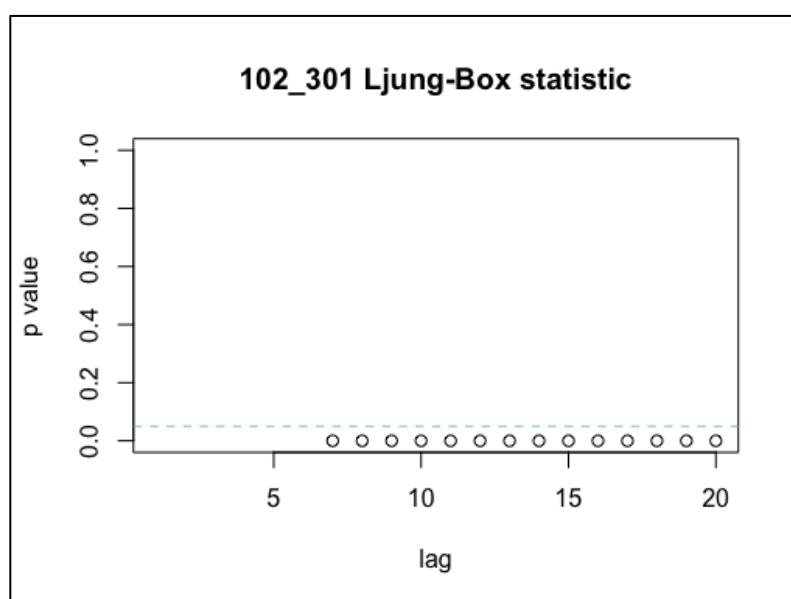


Figure.9 Ljung Box test for SARIMA(1,0,2)x(3,0,1)₁₂

【模型修正】

由於上述的模型都配適不佳，我們考慮將進行差分來 de-trend 並重複上述步驟來進行模型診斷。

五、進行差分來 de-trend 進行 order 選取

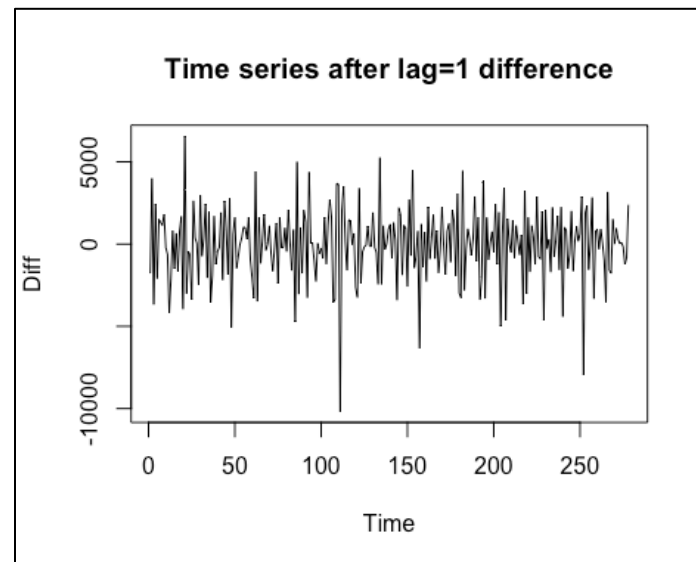


Figure.10 進行一階差分後原始出生人數時間序列圖

從時間序列圖 Figure.10 可以發現，確實進行一階差分後更加的平穩，再進行 ADF 檢定，來檢驗此資料是不是屬於 stationary，以下為檢定進行一階差分後資料的結果：

Augmented Dickey-Fuller Test (diff1)

data: data_diff1

Dickey-Fuller = -9.1596, Lag order = 6, p-value = 0.01

alternative hypothesis: stationary

Warning message:

In adf.test(data_diff1) : p-value smaller than printed p-value

根據檢定結果，在給定 $\alpha=0.05$ 的情形下，我們可以拒絕虛無假設並宣稱此筆進行一階差分後的資料屬於 stationary，而實際 p-value 甚至也是小於 0.01。此外，常理來說應該先做季節性差分早於一般差分，但根據觀察季節性遞減速度的還算快。因此，我們決定先不進行季節性的差分，只保留進行一般差分。

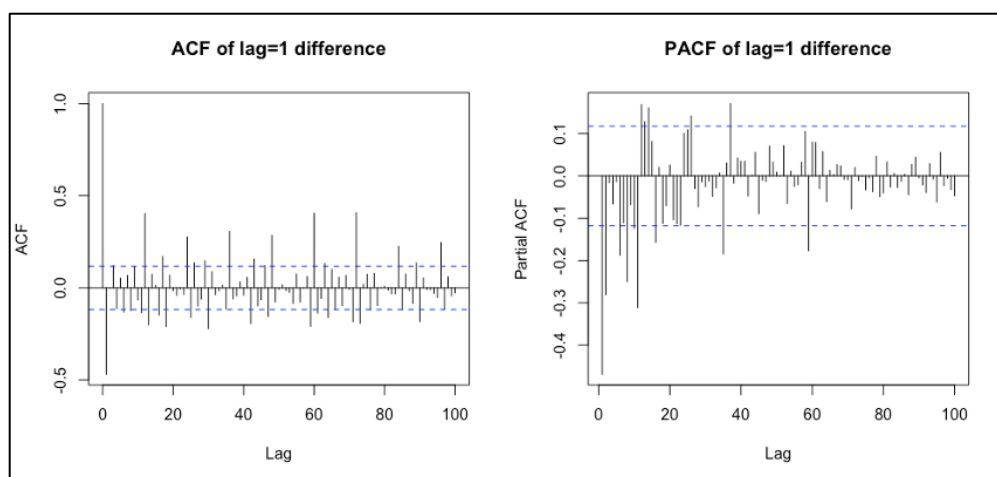


Figure.11 進行一階差分後原始出生人數 ACF、PACF 圖

【Order 選取與參數估計】

判斷：具有週期為 12 的季節性、ACF tails off、PACF cuts off at lag=3 or 5
ACF 前後有 1 期顯著，非季節性部分則考慮(p,d)=(1,0), (1,1), (0,1)

Table.4 模型選取與其相對應參數

	Ar1	Ma1	Sar1	Sar2	Sar3	Sar4	Sar5
SARIMA(0,1,0)x(3,0,0) ₁₂			0.3187	0.0686	0.2369		
s.e			0.0581	0.0635	0.0619		
SARIMA(0,1,0)x(5,0,0) ₁₂			0.2323	-0.0018	0.1426	0.0698	0.3408
s.e			0.0565	0.0592	0.0608	0.0617	0.0611
SARIMA(0,1,0)x(2,0,0) ₁₂			0.3495	0.1522			
s.e.			0.0591	0.0612			
SARIMA(0,1,1)x(5,0,0) ₁₂		-0.6895	0.1873	0.0234	0.2200	0.0983	0.3295
s.e.		0.0409	0.0583	0.0596	0.0595	0.0610	0.0606
SARIMA(1,1,0)x(5,0,0) ₁₂	-0.5391		0.2425	-0.0387	0.2429	0.0244	0.3607
s.e.	0.0510		0.0562	0.0588	0.0586	0.0603	0.0585
SARIMA(0,0,0)x(0,0,0) ₁₂							
s.e.							

從 Table.4 可以看見在正負 2 個標準差的情形下，Seasonal AR=5 才有明顯異於 0 的情況，故我們優先考慮 Seasonal AR=5，並嘗試將 Seasonal AR=3-1 取 2 試試看，但似乎更不好。

【模型診斷- SARIMA(0,1,1)x(5,0,0)₁₂】

method.1 檢驗殘差是否符合常態

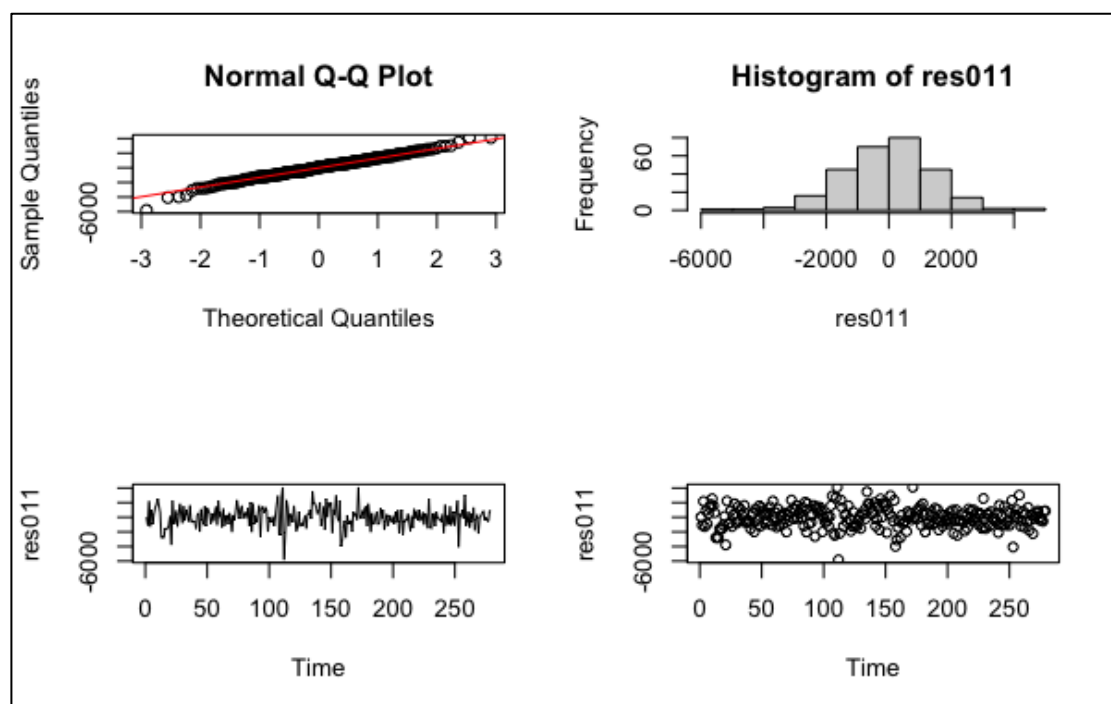


Figure.12 SARIMA(0,1,1)x(5,0,0)₁₂ 殘差常態檢驗

由 Figure.12 殘差檢驗，各方面判定其殘差應是在 0 附近跳動、直方圖也是左右對稱、由 Q-Qplot 判定應也符合常態沒問題。但為求謹慎，我們決定再進行 Shapiro-Wilk normality test，此檢定虛無假設為資料來自於常態，而對立假設則為資料不來自於常態，檢定結果如下：

Shapiro-Wilk normality test

data: res011

W = 0.99085, p-value = 0.07912

根據檢定結果，在 $\alpha=0.05$ 下，我們會不 reject H_0 ，不拒絕虛無假設：資料來自於常態，宣稱殘差是符合常態分配的。

method.2 檢驗殘差是否相關

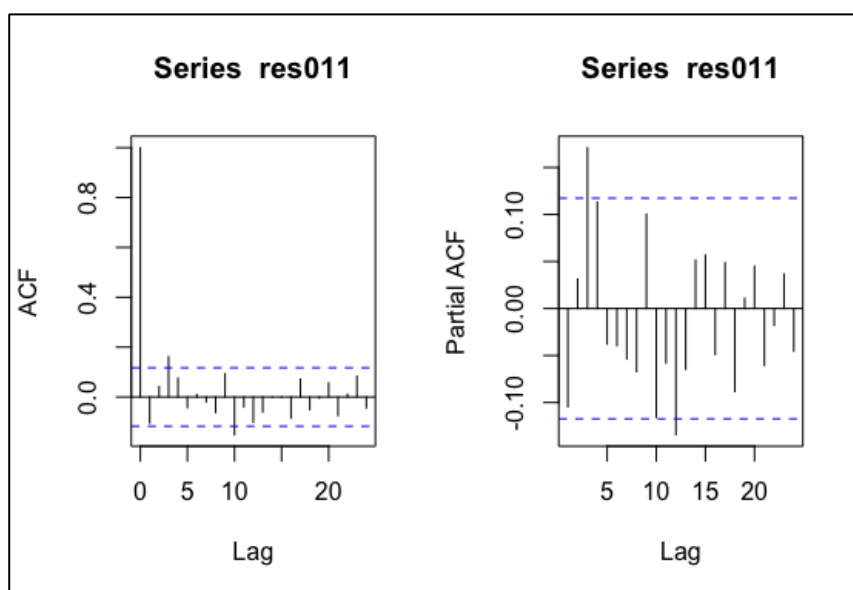


Figure.13 SARIMA(0,1,1)x(5,0,0)₁₂ 殘差 ACF 與 PACF 圖

根據 Figure.13 殘差的 ACF 與 PACF 圖，殘差大抵是在顯著線以下，顯示殘差間大致上無相關性，所以假設誤差是 white noise 是可以被接受的。

method.3 檢驗模型配飾是否適合

接著使用 Ljung Box test 來檢驗模型是否配飾得宜，此檢定

H_0 : the fitted model is adequate H_1 : the fitted model is not adequate

由 Figure.14 所示 Ljung Box test 均無通過，此模型並不是一個配適良好的選擇。

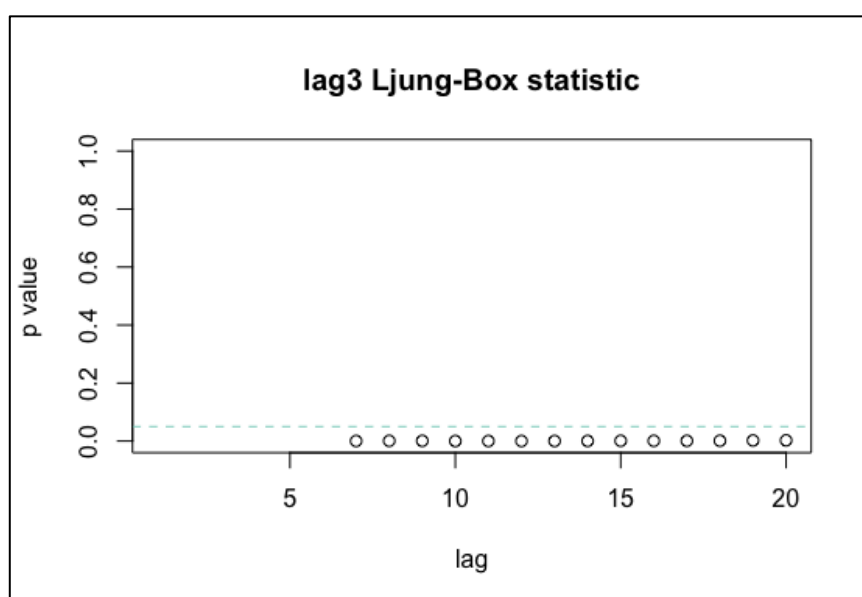


Figure.14 Ljung Box test for SARIMA(0,1,1)x(5,0,0)₁₂

【模型診斷- SARIMA(1,1,0)x(5,0,0)₁₂】

method.1 檢驗殘差是否符合常態

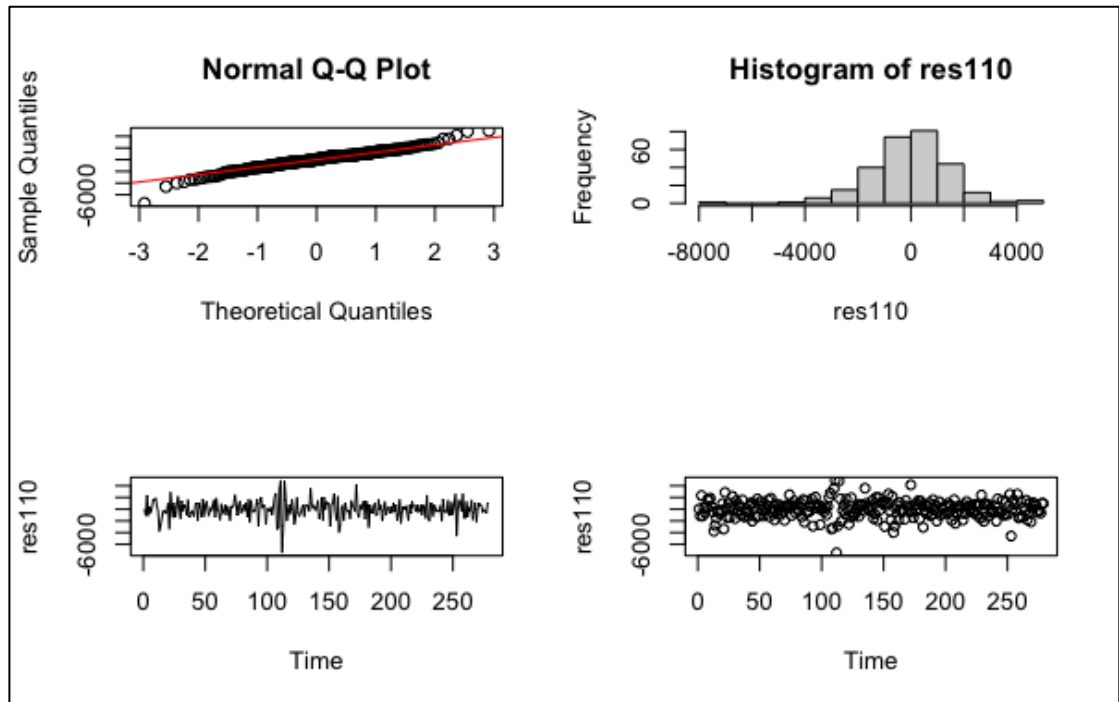


Figure.15 SARIMA(1,1,0)x(5,0,0)₁₂ 殘差常態檢驗

由 Figure.15 殘差檢驗，各方面判定其殘差應是在 0 附近跳動、但直方圖似乎無左右對稱、由 Q-Q plot 判定應也不符合常態。但為求謹慎，我們決定再進行 Shapiro-Wilk normality test，此檢定虛無假設為資料來自於常態，而對立假設則為資料不來自於常態，檢定結果如下：

Shapiro-Wilk normality test

data: res110

W = 0.97337, p-value = 4.616e-05

根據檢定結果，在 $\alpha=0.05$ 下，我們會 reject H_0 ，拒絕虛無假設：資料來自於常態，宣稱殘差是符合不常態分配的。

method.2 檢驗殘差是否相關

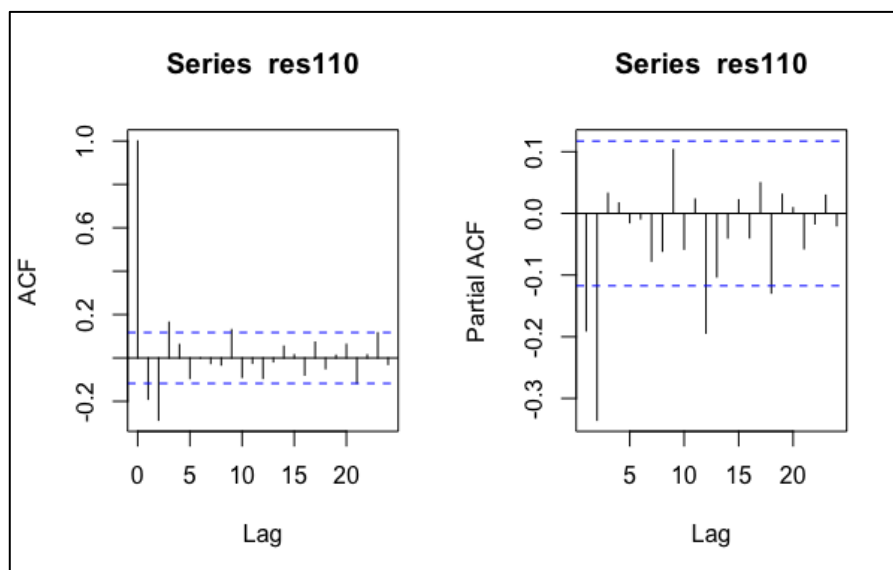


Figure.16 SARIMA(1,1,0)x(5,0,0)₁₂ 殘差 ACF 與 PACF 圖

根據 Figure.16 殘差的 ACF 與 PACF 圖，殘差偶有在顯著線以下，顯示殘差間大致上相關性，假設誤差是 white noise 是不可以被接受的。

method.3 檢驗模型配飾是否適合

接著使用 Ljung Box test 來檢驗模型是否配飾得宜，此檢定

H_0 : the fitted model is adequate H_1 : the fitted model is not adequate

由 Figure.17 所示 Ljung Box test 均無通過，此模型並不是一個配適良好的選擇。

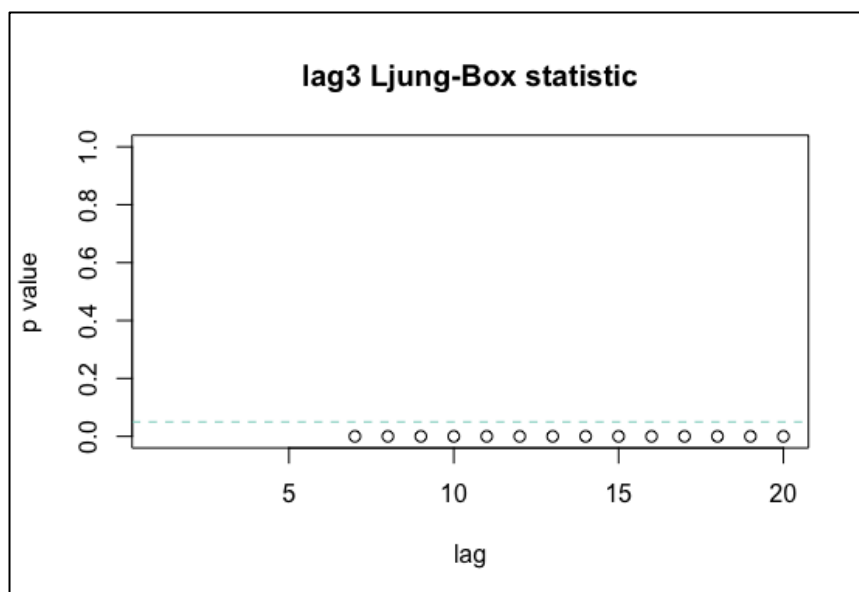


Figure.17 Ljung Box test for SARIMA(1,1,0)x(5,0,0)₁₂

【模型修正】

由於上述的模型也都配適不佳，我們考慮將 Seasonal AR order 選取小一點，並嘗試季節性差分。

六、使用差分來 de-trend 進行 order 選取（含季節性差分）

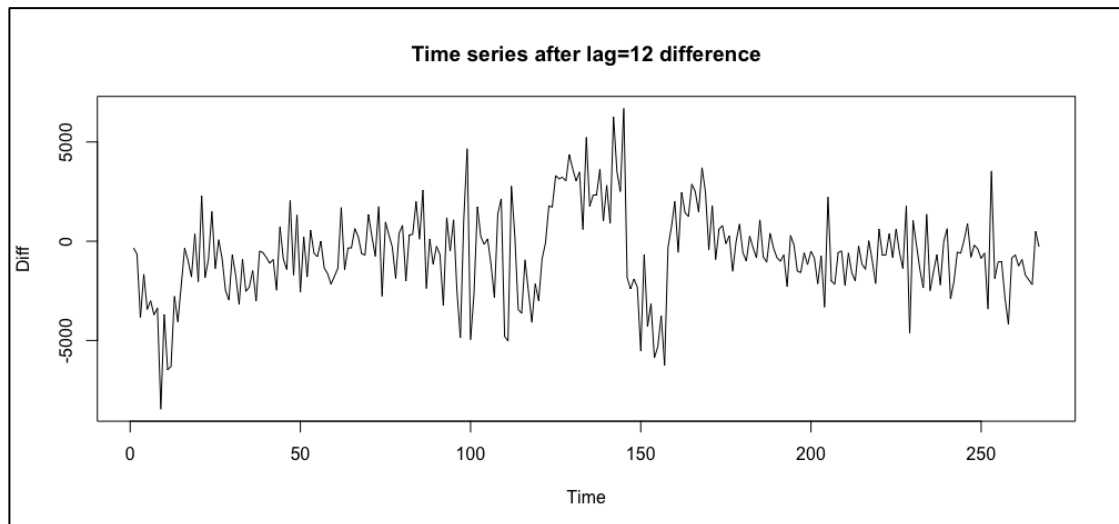


Figure.18 進行季節性差分後出生人數時間序列圖

從時間序列圖 Figure.18 可以發現，確實進行季節性差分後比原始時間序列的平穩，再進行 ADF 檢定，來檢驗此資料是不是屬於 stationary，以下為檢定進行一階差分後資料的結果：

Augmented Dickey-Fuller Test

```
data: data_diff12
```

```
Dickey-Fuller = -4.4091, Lag order = 6, p-value = 0.01
```

```
alternative hypothesis: stationary
```

```
Warning message:
```

```
In adf.test(data_diff12) : p-value smaller than printed p-value
```

根據檢定結果，在給定 $\alpha=0.05$ 的情形下，我們可以拒絕虛無假設並宣稱此筆進行一階差分後的資料屬於 stationary，而實際 p-value 甚至也是小於 0.01。

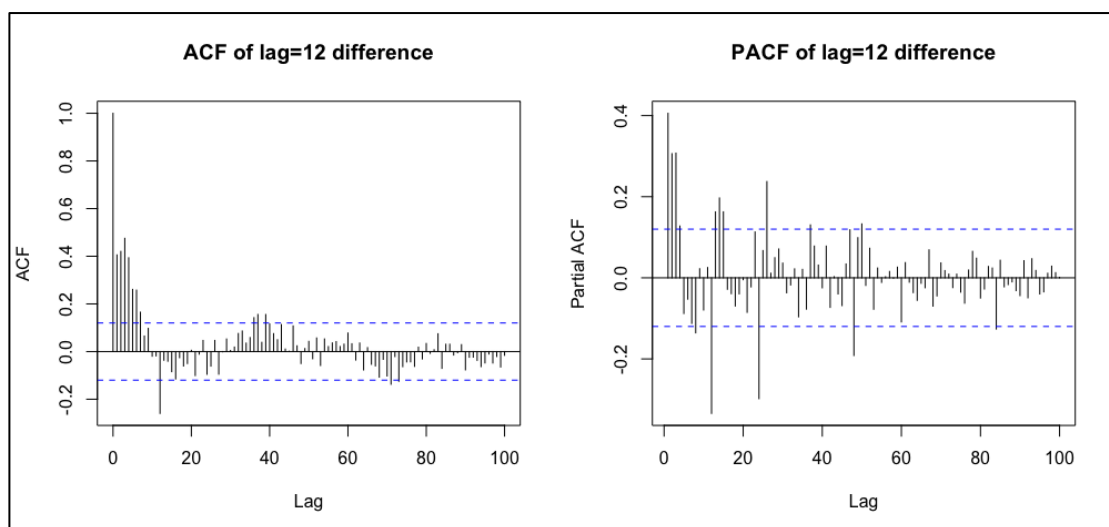


Figure.19 進行季節性差分後出生人數 ACF、PACF 圖

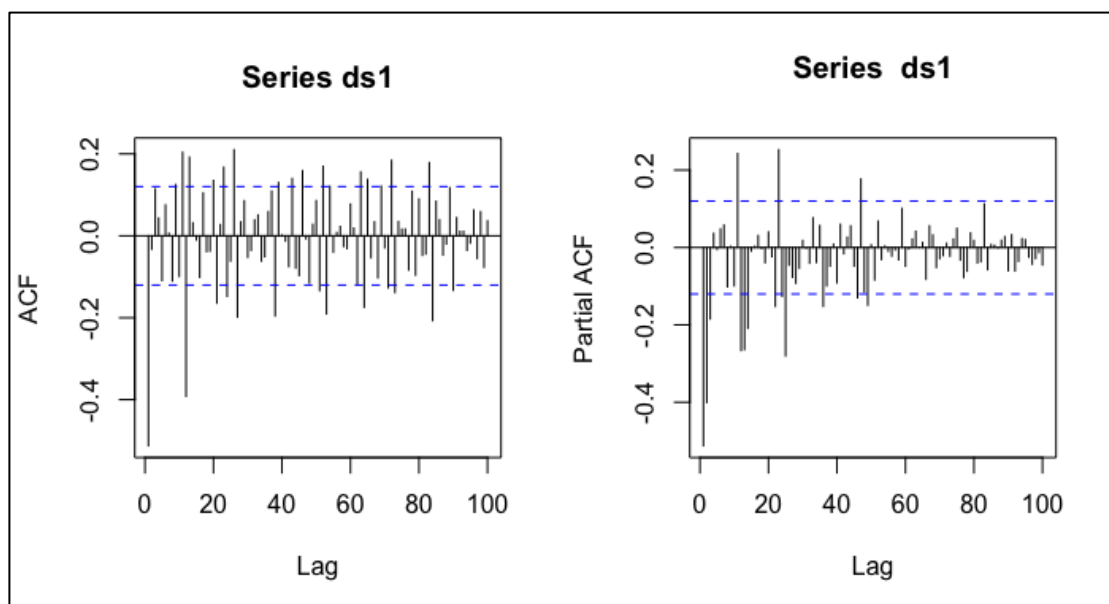


Figure.20 進行季節性與一般差分後出生人數 ACF、PACF 圖

判斷：具有週期為 12 的季節性、ACF tails off、PACF cuts off at lag=3 or 4
ACF 前後有顯著，非季節性部分則考慮(p,d)不超過 3 的 order。

Table.5 模型選取與其相對應參數

	Ar1	Ma1	Sar1	Sar2	Sar3	Sar4	Sma1
SARIMA(0,0,0)x(3,1,1) ₁₂			0.3959	0.0996	0.2435		-0.6290
s.e			0.1549	0.0748	0.0682		0.1547
SARIMA(0,0,0)x(4,1,1) ₁₂			0.3357	0.0890	0.2376	0.0376	-0.5742
s.e			0.1951	0.0767	0.0690	0.0854	0.1865

	AR 最高	MA 最高	Sar1	Sar2	Sar3	Sar4	Sma1
SARIMA(2,1,1)x(3,1,1) ₁₂	-0.3188	-0.2728	-0.0975	-0.1371	-0.0289		-0.8827
s.e.	0.0853	0.1157	0.0824	0.0763	0.0760		0.0731
SARIMA(3,1,1)x(3,1,1) ₁₂	-0.1920	0.1080	-0.0952	-0.1404	-0.0354		-0.8823
s.e.	0.1927	0.4571	0.0824	0.0763	0.0760		0.0729
SARIMA(2,1,1)x(4,1,1) ₁₂	-0.2861	-0.2961	-0.2151	-0.2651	-0.1410	-0.1816	-0.7503
s.e.	0.0889	0.1173	0.1443	0.1308	0.1157	0.0927	0.1600
SARIMA(3,1,1)x(4,1,1) ₁₂	-0.2204	0.1355	-0.2312	-0.2873	-0.1646	-0.2023	-0.7275
s.e.	0.1451	0.3639	0.1585	0.1426	0.1250	0.0981	0.1797
SARIMA(0,1,2)x(4,1,1) ₁₂		-0.2567	-0.2289	-0.3137	-0.1542	-0.2163	-0.7193
		0.0691	0.1503	0.1315	0.1175	0.0915	0.1690
SARIMA(0,1,3)x(4,1,1) ₁₂		0.0938	-0.2432	-0.3211	-0.1778	-0.2154	-0.7017
		0.0634	0.1512	0.1313	0.1778	0.0921	0.1684

觀察 Table.5 我們判斷在季節性 AR=3 時，參數估計雖有顯著異於 0，但加上非季節性部分後季節性 AR=3 就不會呈現一個合理的參數估計結果；反之，AR=4 加上非季節性部分最高項係數則有異於 0，因此挑選了幾個模型符合參數估計準則來去進行配飾及診斷。

【模型診斷- SARIMA(2,1,1)x(4,1,1)₁₂】

method.1 檢驗殘差是否符合常態

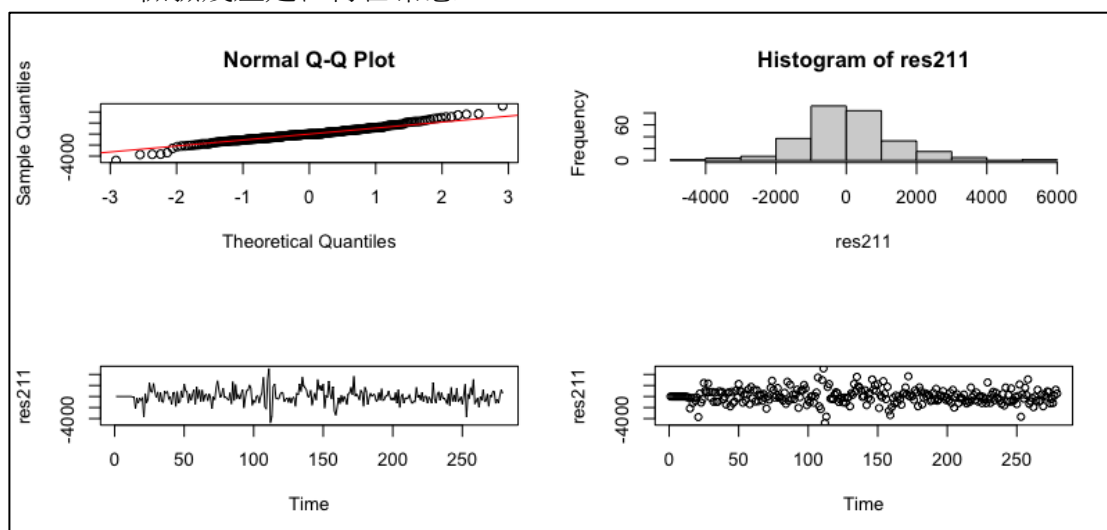


Figure.21 SARIMA(2,1,1)x(4,1,1)₁₂ 殘差常態檢驗

由 Figure.21 殘差檢驗，各方面判定其殘差應是在 0 附近跳動、直方圖有點左右不對稱、由 Q-Qplot 判定應不太符合常態。但為求謹慎，我們決定再進行 Shapiro-Wilk normality test，此檢定虛無假設為資料來自於常態，而對

立假設則為資料不來自於常態，檢定結果如下：

Shapiro-Wilk normality test

data: res211

W = 0.97683, p-value = 0.0001688

根據檢定結果，在 $\alpha=0.05$ 下，我們會 reject H_0 ，拒絕虛無假設：資料來自於常態，宣稱殘差是不符合常態分配的。

method.2 檢驗殘差是否相關

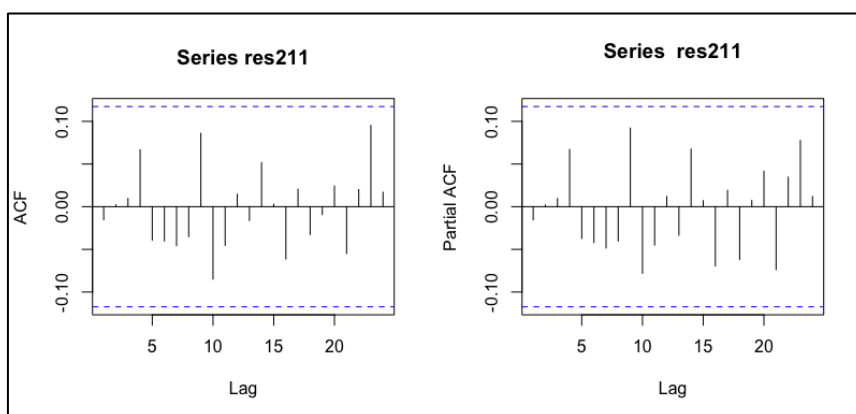


Figure.22 SARIMA(2,1,1)x(4,1,1)₁₂ 殘差 ACF 與 PACF 圖

根據 Figure.22 殘差的 ACF 與 PACF 圖，殘差全部都是在顯著線以下，顯示殘差間大致上無相關性，所以假設誤差是 white noise 是可以被接受的。

method.3 檢驗模型配飾是否適合

接著使用 Ljung Box test 來檢驗模型是否配飾得宜，此檢定

H_0 : the fitted model is adequate H_1 : the fitted model is not adequate

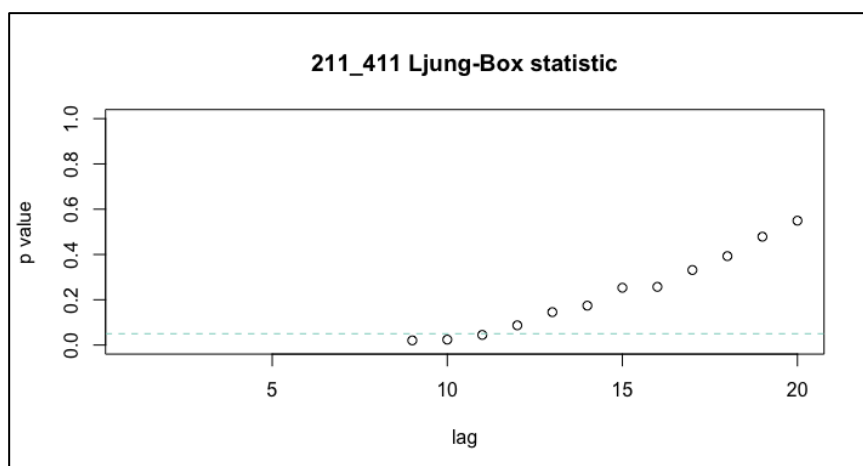


Figure.23 Ljung Box test for SARIMA(2,1,1)x(4,1,1)₁₂

由 Figure.23 所示 Ljung Box test 僅少部分未通過，此模型還算配適良好的選擇。

【模型診斷- SARIMA(3,1,1)x(4,1,1)₁₂】

method.1 檢驗殘差是否符合常態

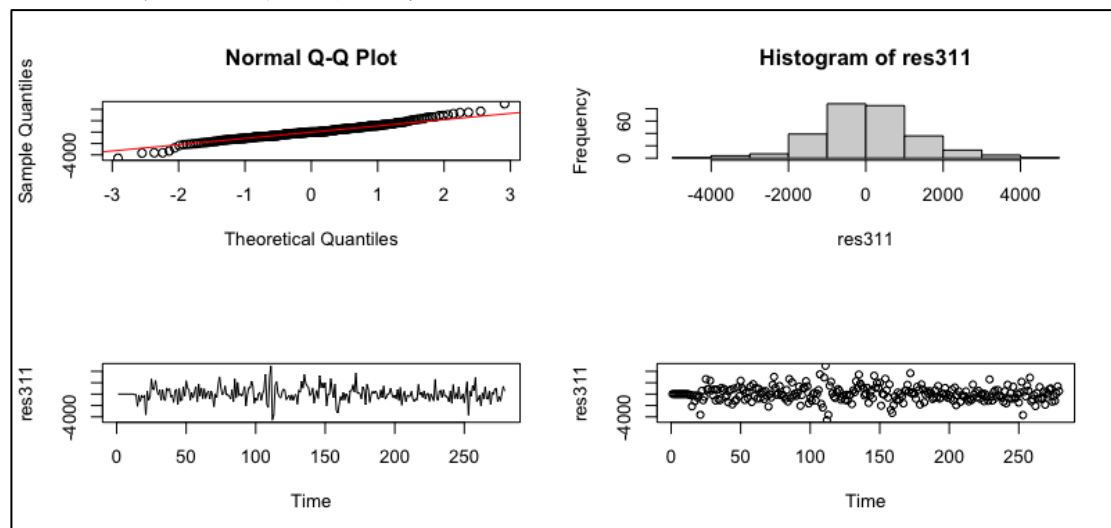


Figure.23 SARIMA(3,1,1)x(4,1,1)₁₂ 殘差常態檢驗

由 Figure.23 殘差檢驗，各方面判定其殘差應是在 0 附近跳動、直方圖有點左右還算對稱、由 Q-Qplot 判定應不太符合常態。但為求謹慎，我們決定再進行 Shapiro-Wilk normality test，此檢定虛無假設為資料來自於常態，而對立假設則為資料不來自於常態，檢定結果如下：

Shapiro-Wilk normality test

data: res311

W = 0.97959, p-value = 0.0005045

根據檢定結果，在 $\alpha=0.05$ 下，我們會 reject H_0 ，拒絕虛無假設：資料來自於常態，宣稱殘差是不符合常態分配的。

method.2 檢驗殘差是否相關

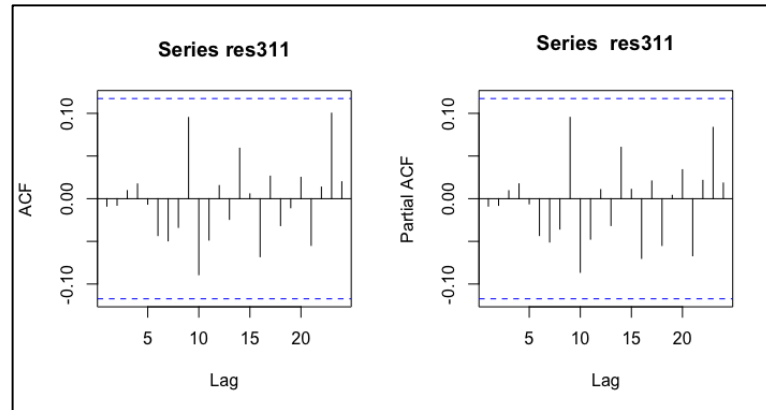


Figure.25 SARIMA(3,1,1)x(4,1,1)₁₂ 殘差 ACF 與 PACF 圖

根據 Figure.25 殘差的 ACF 與 PACF 圖，殘差全部都是在顯著線以下，顯示殘差間大致上無相關性，所以假設誤差是 white noise 是可以被接受的。

method.3 檢驗模型配飾是否適合

接著使用 Ljung Box test 來檢驗模型是否配飾得宜，此檢定

H_0 : the fitted model is adequate H_1 : the fitted model is not adequate

由 Figure.26 所示 Ljung Box test 僅少部分未通過，此模型還算配適良好的選擇。

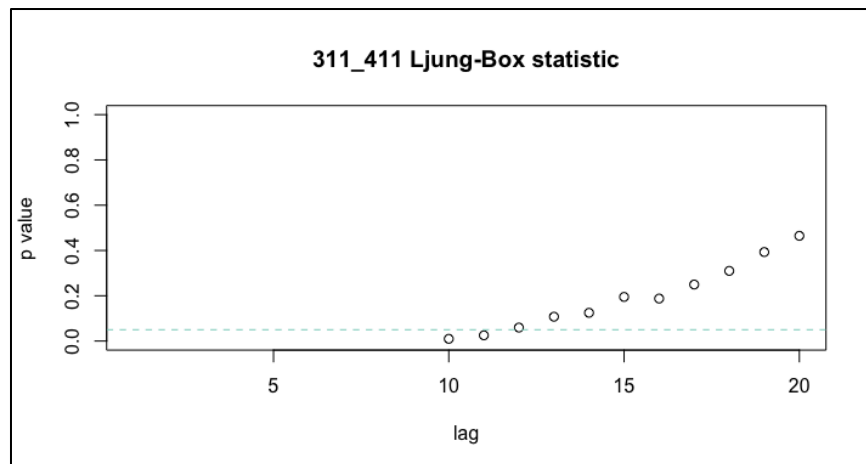


Figure.26 Ljung Box test for SARIMA(2,1,1)x(4,1,1)₁₂

【模型診斷- SARIMA(0,1,2)x(4,1,1)₁₂】

method.1 檢驗殘差是否符合常態

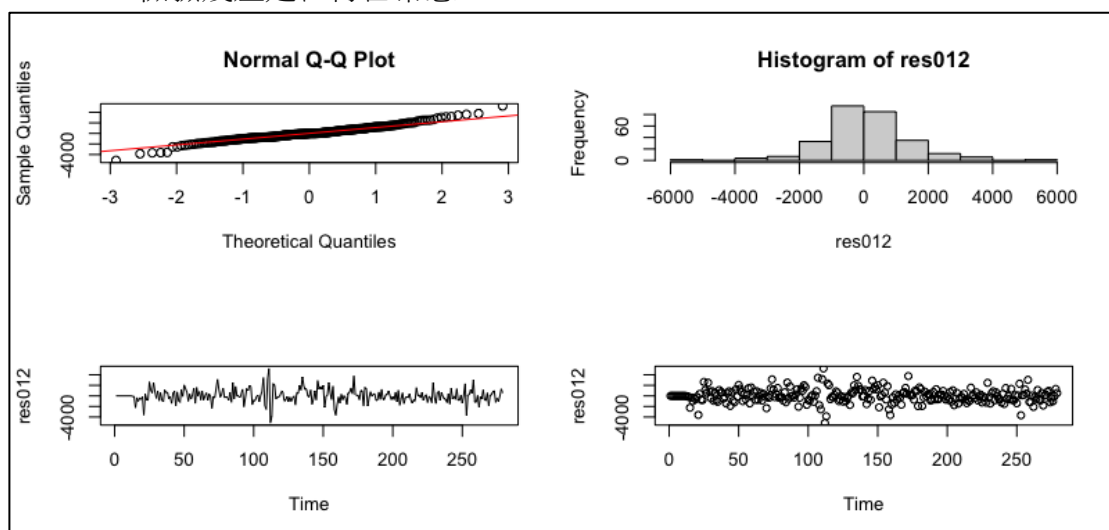


Figure.27 SARIMA(0,1,2)x(4,1,1)₁₂ 殘差常態檢驗

由 Figure.27 殘差檢驗，各方面判定其殘差應沒有都在 0 附近跳動、直方圖有點左右還算對稱、由 Q-Qplot 判定應不太符合常態。但為求謹慎，我們決定再進行 Shapiro-Wilk normality test，此檢定虛無假設為資料來自於常態，而對立假設則為資料不來自於常態，檢定結果如下：

Shapiro-Wilk normality test

data: res012

W = 0.97411, p-value = 6.041e-05

根據檢定結果，在 $\alpha=0.05$ 下，我們會 reject H_0 ，拒絕虛無假設：資料來自於常態，宣稱殘差是不符合常態分配的。

method.2 檢驗殘差是否相關

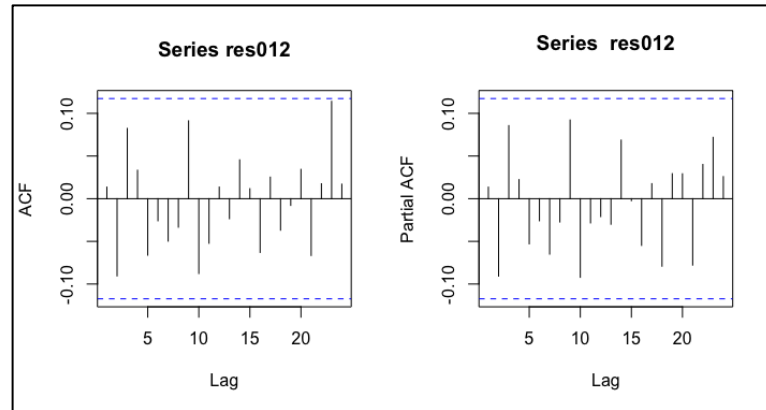


Figure.28 SARIMA(0,1,2)x(4,1,1)₁₂ 殘差 ACF 與 PACF 圖

根據 Figure.28 殘差的 ACF 與 PACF 圖，殘差全部都是在顯著線以下，顯示殘差間大致上無相關性，所以假設誤差是 white noise 是可以被接受的。

method.3 檢驗模型配飾是否適合

接著使用 Ljung Box test 來檢驗模型是否配飾得宜，此檢定

H_0 : the fitted model is adequate H_1 : the fitted model is not adequate

由 Figure.29 所示 Ljung Box test 有部分未通過，因此我們認為此模型配適尚差強人意。

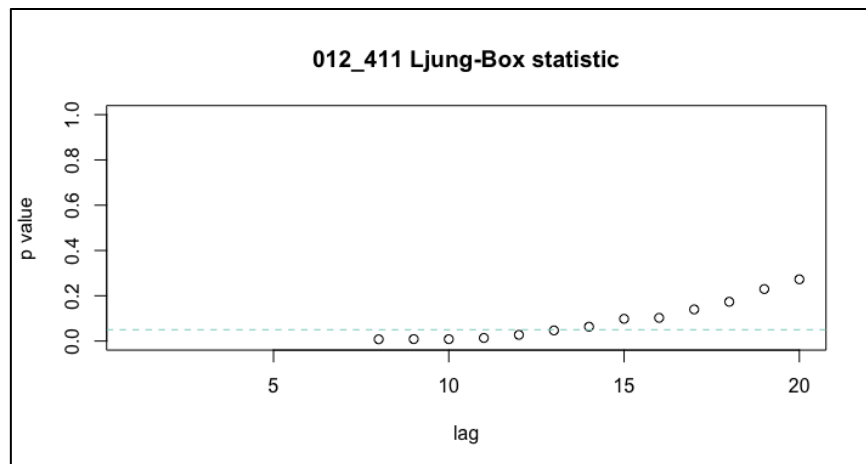


Figure.29 Ljung Box test for SARIMA(0,1,2)x(4,1,1)₁₂

【模型診斷- SARIMA(0,1,3)x(4,1,1)₁₂】

method.1 檢驗殘差是否符合常態

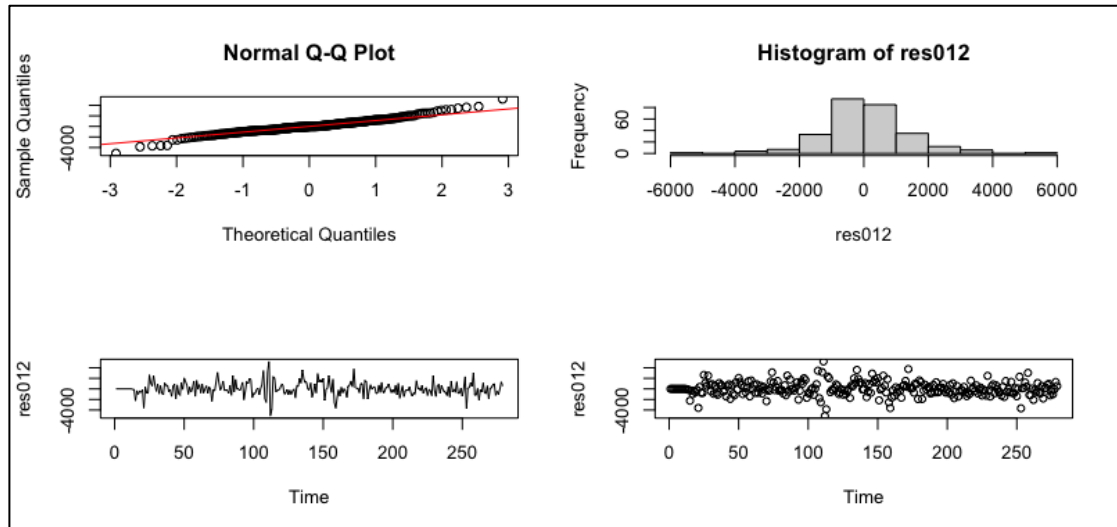


Figure.30 SARIMA(0,1,3)x(4,1,1)₁₂ 殘差常態檢驗

由 Figure.30 殘差檢驗，各方面判定其殘差應沒有都在 0 附近跳動、直方圖有點左右還算對稱、由 Q-Qplot 判定應不太符合常態。但為求謹慎，我們決定再進行 Shapiro-Wilk normality test，此檢定虛無假設為資料來自於常態，而對立假設則為資料不來自於常態，檢定結果如下：

Shapiro-Wilk normality test

data: res013

W = 0.97702, p-value = 0.0001816

根據檢定結果，在 $\alpha=0.05$ 下，我們會 reject H_0 ，拒絕虛無假設：資料來自於常態，宣稱殘差是不符合常態分配的。

method.2 檢驗殘差是否相關

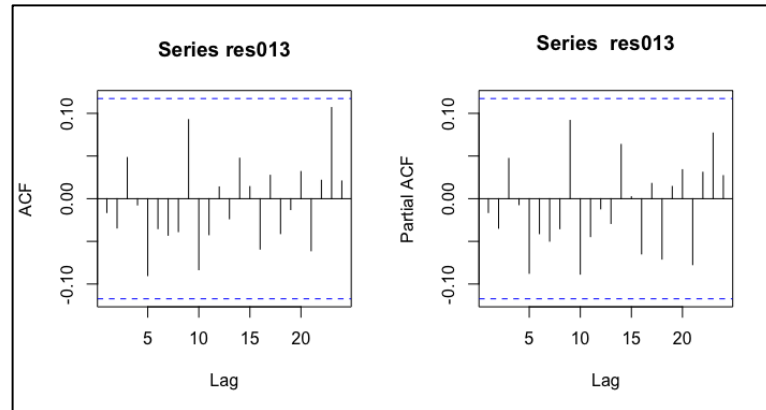


Figure.31 SARIMA(0,1,3)x(4,1,1)₁₂ 殘差 ACF 與 PACF 圖

根據 Figure.31 殘差的 ACF 與 PACF 圖，殘差全部都是在顯著線以下，顯示殘差間大致上無相關性，所以假設誤差是 white noise 是可以被接受的。

method.3 檢驗模型配飾是否適合

接著使用 Ljung Box test 來檢驗模型是否配飾得宜，此檢定

H_0 : the fitted model is adequate H_1 : the fitted model is not adequate

由 Figure.31 所示 Ljung Box test 有部分未通過，因此此模型配適上應算普通。

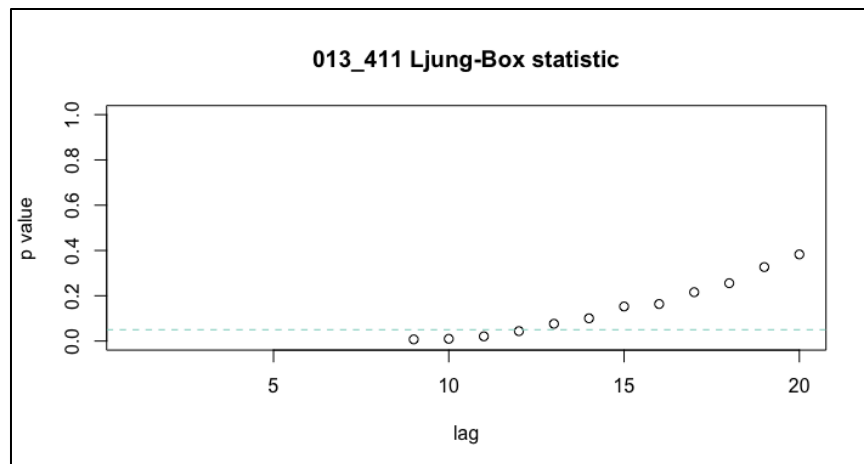


Figure.32 Ljung Box test for SARIMA(0,1,3)x(4,1,1)₁₂

七、模型選擇

Table.6 模型選取比較

	Log likelihood	AIC	Ljung Box
SARIMA(2,1,1)x(4,1,1) ₁₂	-2300.54	4617.08	還算良好
SARIMA(3,1,1)x(4,1,1) ₁₂	-2299.71	4617.41	還算良好
SARIMA(0,1,2)x(4,1,1) ₁₂	-2302.27	4618.55	差強人意
SARIMA(0,1,3)x(4,1,1) ₁₂	-2301.17	4618.34	普通

模型：

八、預測

在建立模型時，我們保留最後 12 期（2022 年 4 月到 2023 年 3 月）出生資料來驗證與們的預測值與實際值差異於 Table.7，並繪製 12-step 預測的時間序列圖 Figure.33。

Table.7 模型選取比較

	實際值	預測值	95%預測下界	95%預測上界	預測誤差	時間
1-step	11222	10809.558	8184.229	13434.888	-412.442	111/4
2-step	9442	11137.560	8466.642	13808.477	1695.560	5
3-step	10943	10850.156	8052.238	13648.074	-92.844	6
4-step	10950	11323.972	8247.441	14400.503	373.972	7
5-step	11902	11595.533	8376.395	14814.672	-306.467	8
6-step	12217	11573.094	8205.258	14940.929	-643.906	9
7-step	12242	13450.769	9918.629	16982.909	1208.769	10
8-step	12348	12802.435	9127.792	16477.077	454.435	11
9-step	12178	13542.02	9728.932	17355.107	1364.020	12
10-step	10948	11574.504	7624.573	15524.436	626.504	112/1
11-step	10113	9679.715	5599.513	13759.917	-433.285	2
12-step	12506	12626.191	8419.642	16832.739	120.191	3

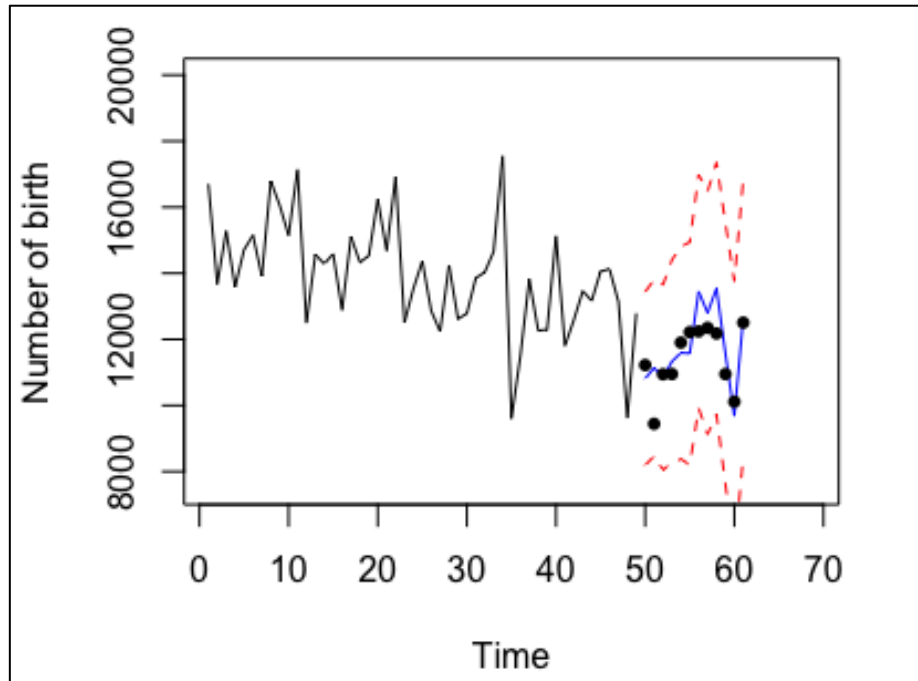


Figure.33 2022 年 4 月到 2023 年 3 月時間序列預測值實際觀測值比較

Figure.33 藍線表示預測值、紅線分別代表 95%預測區間上下界、黑色圓點則代表實際值有便於對照。可以從表格及比較圖表發現所有的預測值都有落在預測區間，唯獨在 2022 年 5 月有明顯落差。

九、結論

根據我們的預測模型可以看見其容易在極值的時候失準，例如：在出生大月的 10、12 月、而對於 2022 年 5 月的過於樂觀預估細究可能原因為，該月份為國家景氣亮紅燈的時期，而後續月份的回升推測為在 2021 年 7 月以後政府的防疫警戒降為 2 級，同時提供眾多紓困補助來振興經濟。

無論是經由原始資料、預測資料我們都可以得知台灣出生人口與社會經濟、政令的影響可以說是息息相關。給予我們內政服可參酌的施政方向為，若想解決逐漸成為台灣高齡化社會的問題，對於父母經濟補貼政策段期內是有效的沒錯；但長遠之計的策略在於

- (1) 政府有沒有帶領國家進入好的景氣
- (2) 政府對於好的鼓勵生育時機有沒有把握住

若有掌握此 2 原則，因能對我國的出生人口有很大的幫助。

十、附錄

1. 本次期末研究報告資料來源：

中華民國統計資訊網

<https://nstatdb.dgbas.gov.tw/dgbasall/webMain.aspx?k=dgmain>

中華民國統計資訊網
National Statistics, R.O.C. (Taiwan)

總體統計資料庫

新版 總體統計資料庫

關鍵字

單表查詢 跨表查詢 視圖圖 快捷工具 統計驛站 首頁

出生、死亡統計

統計期: 107 年 1 月至 112 年 4 月
週期: 月(含年)
計算模式: 統計值
輸出格式: 網頁
查詢

領域: 人口統計
資料週期: 年/月
最新公布資料: (年)民國111年、(月)民國112年4月
最早歷史資料: (年)民國80年、(月)民國89年1月
資料來源: 內政部統計處

統計項

指標 (全選)

- ☒ 出生人數(人)
- ☒ 粗出生率(‰)
- ☒ 男性出生人數(人)
- ☒ 女性出生人數(人)
- ☒ 死亡人數(人)
- ☒ 粗死亡率(‰)
- ☒ 男性死亡人數(人)
- ☒ 女性死亡人數(人)
- ☒ 自然增加人數(人)
- ☒ 自然增加率(‰)
- ☒ 社會增加人數(人)
- ☒ 社會增加率(‰)
- ☒ 總增加人數(人)
- ☒ 總增加率(‰)

1 村里鄰戶數暨人口數
2 現住原住民人口數
3 出生、死亡統計
4 戶籍登記現住人口數-按五
5 結婚、離婚統計
6 結婚人數-按雙方國籍分
7 離婚人數-按雙方國籍分
8 嬰兒出生數-按生母國籍別
9 15歲以上人口教育程度
10 婚姻狀況統計
11 平均餘命
12 國際、內戶籍遷入、遷出
13 入出境人數統計-按入境
14 入出境人數統計-按出境
▶ 警政與消防統計
▶ 社會保險及福利統計

附錄 Figure.I 中華民國統計資訊網總體統計資料庫

2. 文獻回顧參考資料

王英傑。2015，中央大學 《景氣循環對婚育行為之影響》，p.8

戴庭玉。2011，陽明交通大學 《美國 2007 年至 2009 年引發環球金融風暴的制度因素》，p.7

3. 台灣社會新聞報導回顧參考資料

羅真、何定照、張佳麗、邱苑苓。2019，udn 新聞網

《這些神秘年分 台灣為何生育率掉底又上升》

<https://theme.udn.com/theme/story/6773/3721938>