

消費者物價指數與失業率 相關性分析

題目：消費者物價指數與失業率之相關性分析

學校：國立成功大學

系級：統計學系 111 級

姓名：蔡儀萱 林少穎

學號：H24076150 H24071273

目錄

一、研究動機	03
二、資料介紹	03
三、建模流程	04
四、模型建置	05
1. 消費者物價指數的模型建置	05
2. 失業率的模型建置	14
五、相關性分析	27
六、結論	28
七、參考資料	28
八、課堂心得與建議	29

一、研究動機

消費者物價指數 (CPI) 在經濟學上，是反映與消費者有關的生活產品和勞務價格的物價變動指標，為通貨膨脹的重要指標之一，若 CPI 指數上漲表示有通貨膨脹的跡象，因此在相同的所得水準下，購買力會隨物價上漲而降低，影響層面相當廣泛，所以政府在做決策時經常參考 CPI 指數。而 CPI 穩定、就業充分及 GDP 增長往往是最重要的社會經濟目標，因此我們想探討在整體社會經濟下，CPI 指數與失業率有無相關，以下將進行消費者物價指數與失業率之相關性分析。

二、資料介紹

1. 消費者物價指數

資料來源：中華民國統計資料網

時間範圍：西元 1981 年 1 月至西元 2020 年 10 月

資料比數：共 478 筆

時間間隔：以月為單位

2. 失業率

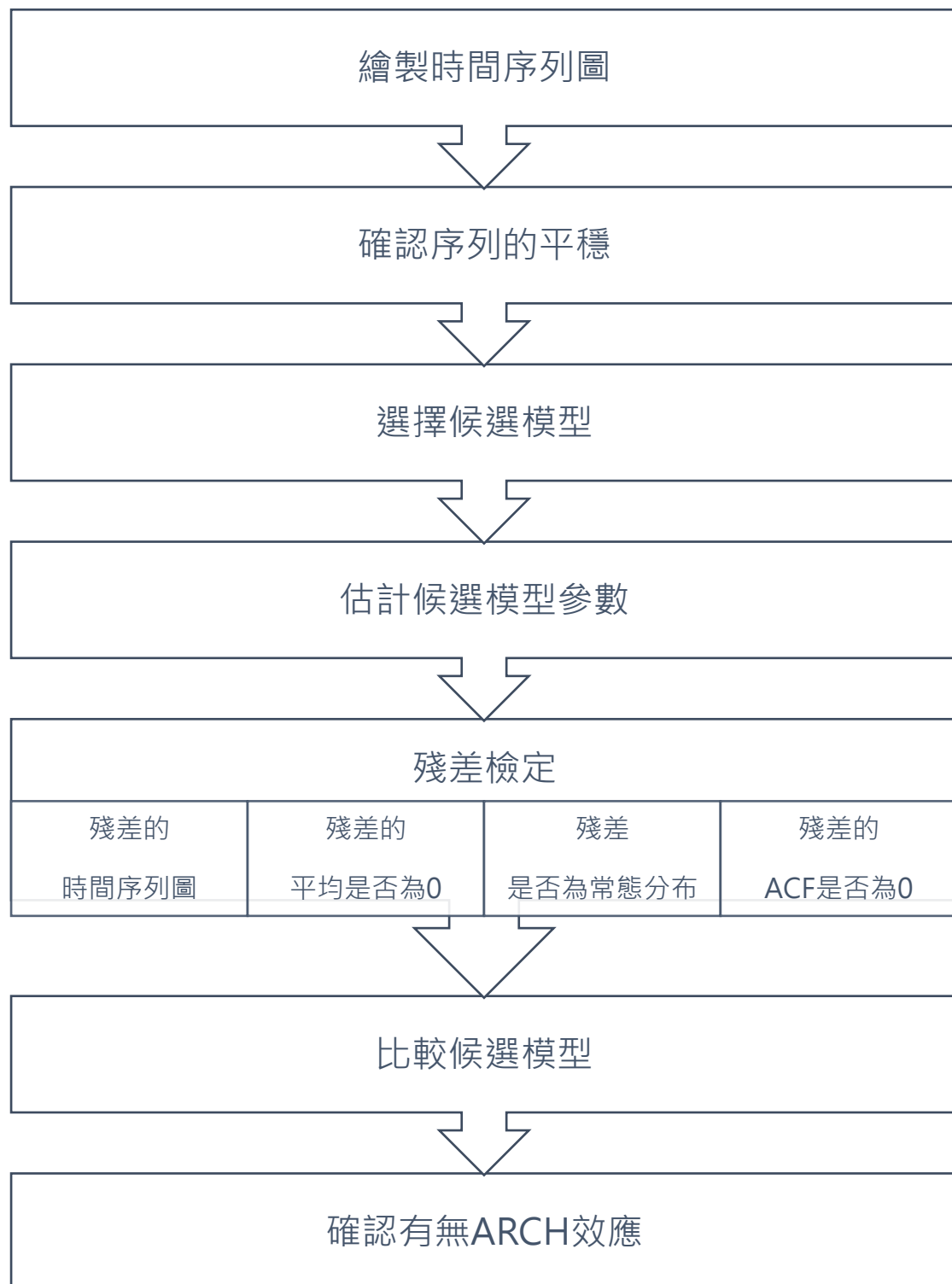
資料來源：中華民國統計資料網

時間範圍：西元 1981 年 1 月至西元 2020 年 10 月

資料比數：共 478 筆

時間間隔：以月為單位

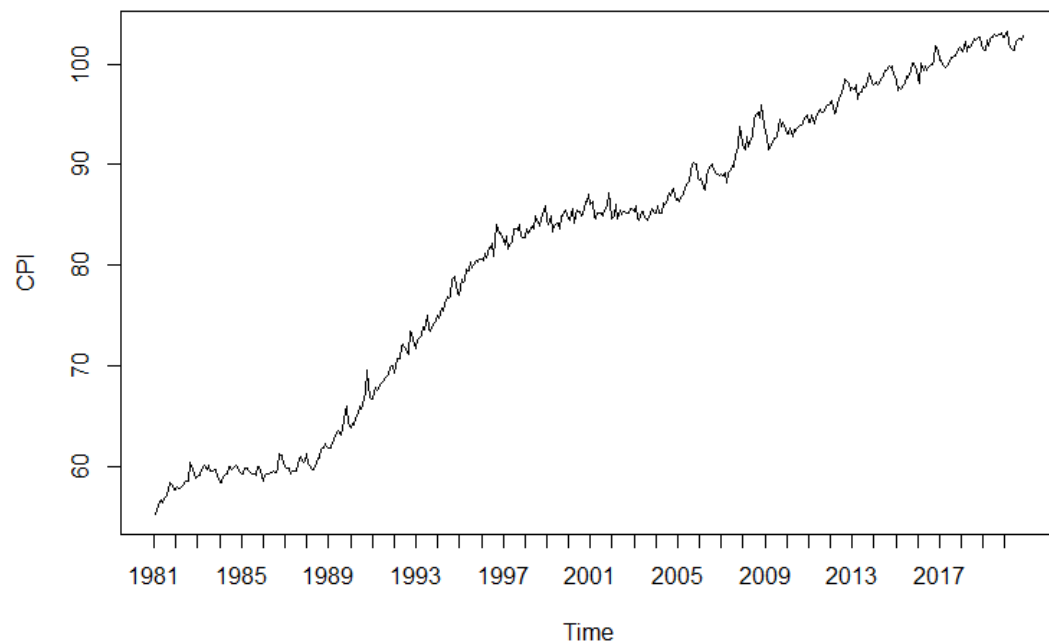
三、建模流程



四、模型建置

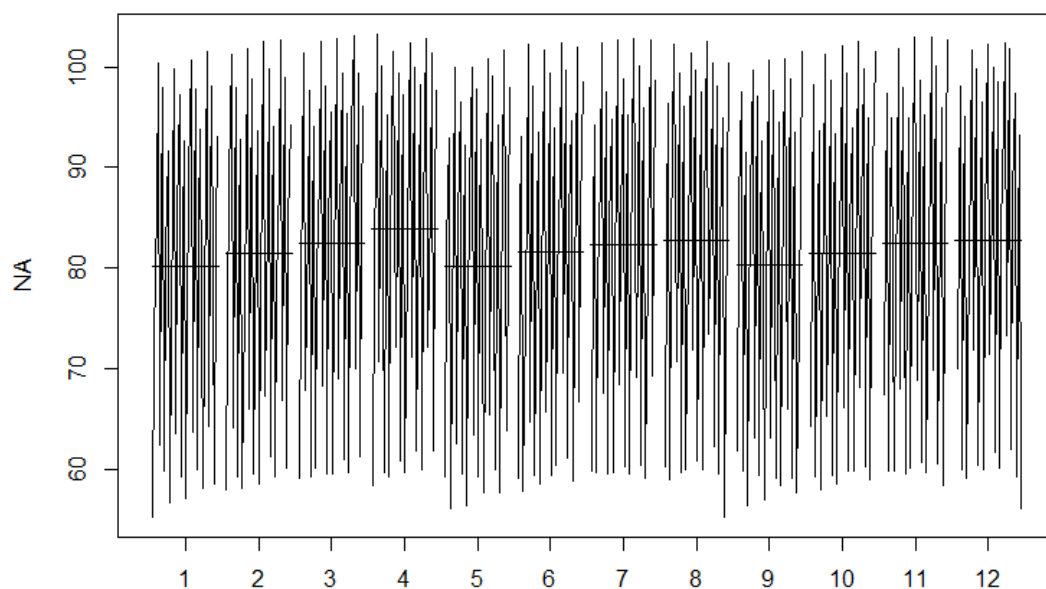
1. 消費者物價指數的模型建置

(1) 繪製消費者物價指數的時間序列圖



由上圖可看出，資料有呈現持續上升的趨勢。

(2) 消費者繪製物價指數的 Month Plot



由上圖可看出，各月份的分布差異並不顯著，因此尚無法判斷此資料是否具有季節效應，需待後續繪製 ACF 及 PACF 圖以利判斷。

(3) 確認資料的平穩

為了確認資料是否需要差分，因此做 Dickey-Fuller Test 及 KPSS Test，以檢定的方式來判斷。

i. Dickey-Fuller Test

H_0 : Non-Stationary. (Need difference.) H_1 : Stationary. (Do not need difference.)	
Dickey-Fuller	-0.7282
p-value	0.9679
Conclusion	Do not reject H_0 , need difference.

ii. KPSS Test

H_0 : Stationary. (Do not need difference.) H_1 : Non-Stationary. (Need difference.)	
KPSS Level	7.8442
p-value	0.01
Conclusion	Reject H_0 , need difference.

從兩個檢定方法都得出資料不平穩，需差分的結論，因此再利用上述兩個檢定判斷是否需要二階差分。

i. Dickey-Fuller Test

H_0 : Non-Stationary. (Need difference.) H_1 : Stationary. (Do not need difference.)	
Dickey-Fuller	-11.153
p-value	0.01
Conclusion	Reject H_0 , don't need difference.

ii. KPSS Test

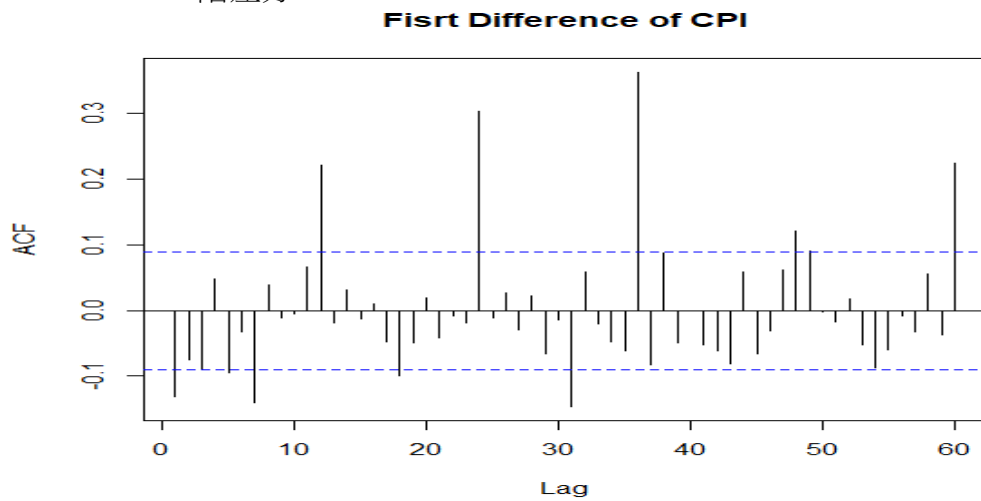
H_0 : Stationary. (Do not need difference.) H_1 : Non-Stationary. (Need difference.)	
KPSS Level	0.14294
p-value	0.1
Conclusion	Do not reject H_0 , don't need difference.

由上述的結果可知此資料一階差分後即平穩，無需做二階差分。

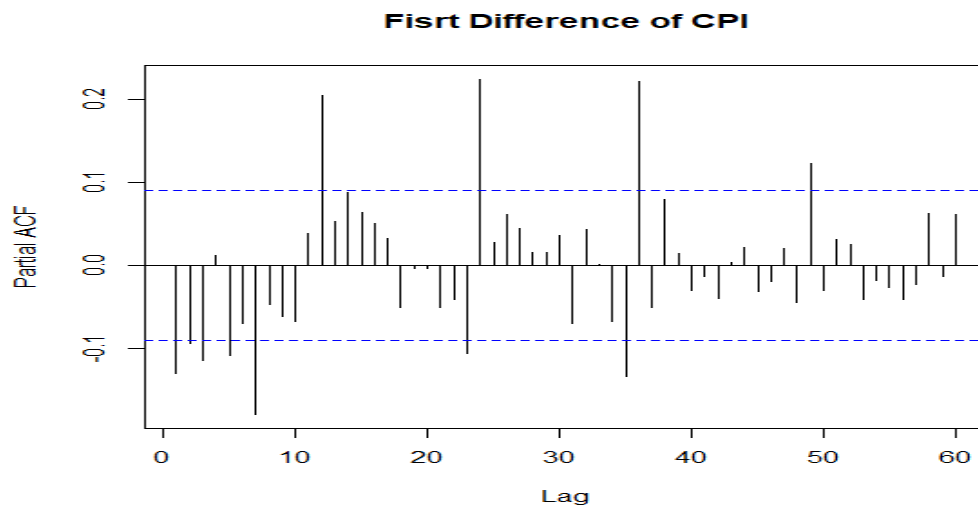
(4) 選擇候選模型

透過觀察資料的 ACF、PACF 及 EACF 圖，選擇合適的候選模型。

i. 一階差分



從 ACF 圖（上圖）可看出，第 7 步時仍有顯著性，且在第 12、24、36、48 及 60 步時也皆超出信賴界線。

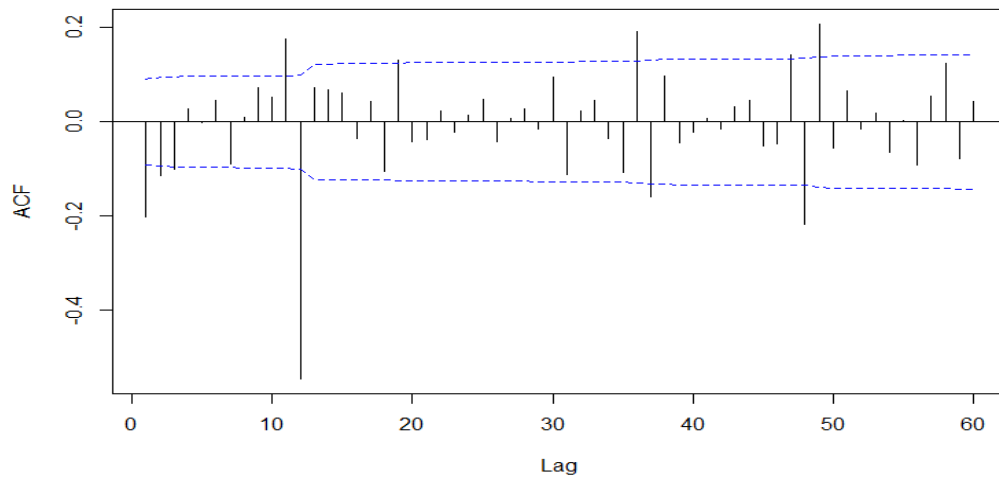


從 PACF 圖（上圖）可看出，第 7 步時仍有顯著性，且在第 12、24、36、48 及 60 步時也皆超出信賴界線。

綜合上述的 ACF 及 PACF 圖可看出，該資料應具有以 12 個月為一單位的季節效應，因此後續將對此資料進行一階差分及季節性一階差分。

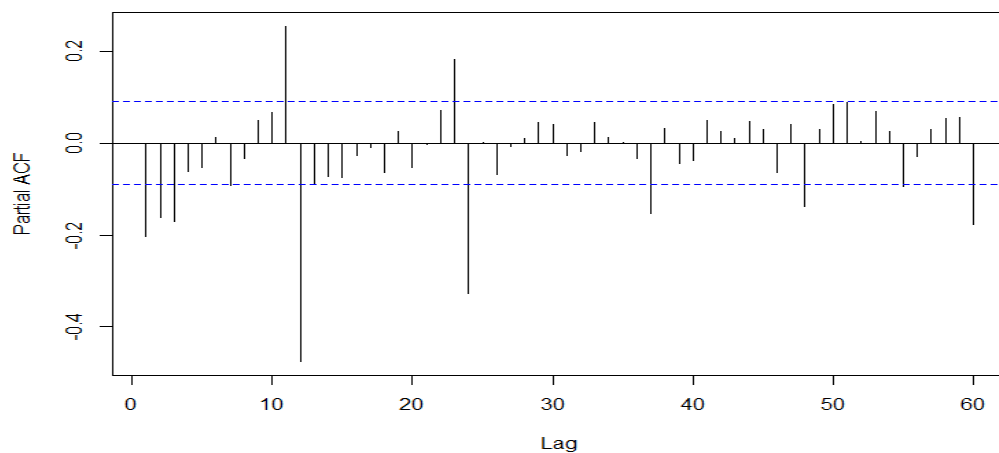
ii. 一階差分與季節性一階差分

Fisrt Difference and Seasonal First Difference of CPI



從 ACF 圖（上圖）可看出，在第 3 步時截斷，而在第 11、12 步後仍有些許步數超出信賴界線，但由於影響資料的步數過大，因此僅挑選 $SARIMA(0,1,3) \times (0,1,1)_{12}$ 候選模型。

Fisrt Difference and Seasonal First Difference of CPI



從 PACF 圖（上圖）可看出，第 7 步時仍有顯著性，且在第 11、12、23、24、36、48 及 60 步時也皆超出信賴界線，由於影響資料的步數過大，因此暫不挑選候選模型。

AR/MA		0	1	2	3	4	5	6	7	8	9	10	11	12	13
0	x	x	x	o	o	o	o	o	o	o	o	x	x	o	o
1	x	o	x	o	o	o	o	o	o	o	o	o	x	x	o
2	x	x	x	o	o	o	o	o	o	o	o	o	x	x	o
3	x	x	x	o	o	o	o	o	o	o	o	o	x	o	x
4	x	x	o	x	o	o	o	o	o	o	o	o	x	o	x
5	x	x	o	x	o	o	o	o	o	o	o	o	x	x	o
6	x	x	x	x	o	x	o	o	o	o	o	o	x	x	o
7	x	x	x	x	x	o	o	o	o	o	o	o	x	o	x

從 EACF 圖（上圖）可得知候選模型為 $SARIMA(1,1,1) \times (0,1,1)_{12}$ ，因為 $ARMA(1,1)$ 的三角區域涵蓋較多「o」。

綜合上述，共得 2 個候選模型，以下將進行模型的參數估計。

(5) 估計候選模型參數

i. **SARIMA(0, 1, 3) × (0, 1, 1)₁₂**

$$(1 - B^{12})(1 - B)Y_t = (1 + \hat{\theta}_1 B + \hat{\theta}_2 B^2 + \hat{\theta}_3 B^3)(1 + \hat{\Theta} B^{12})e_t$$

	$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\theta}_3$	$\hat{\Theta}$
Estimator	-0.2243	-0.1231	-0.0498	-0.9023
Standard Error	0.0465	0.0464	0.0443	0.0287

由上表可知， $\hat{\theta}_3$ 估計值的正負兩倍標準差包含 0，因此應將此候選模型調整為 SARIMA(0,1,2) × (0,1,1)₁₂。

ii. **SARIMA(0, 1, 2) × (0, 1, 1)₁₂**

$$(1 - B^{12})(1 - B)Y_t = (1 + \hat{\theta}_1 B + \hat{\theta}_2 B^2)(1 + \hat{\Theta} B^{12})e_t$$

	$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\Theta}$
Estimator	-0.2325	-0.1335	-0.9035
Standard Error	0.0455	0.0450	0.0283

由上表可知，所有參數估計值的正負兩倍標準差皆不包含 0，因此接著對該模型進行離群值檢測。

	IO[1]	IO[2]	IO[3]
n th	188	242	322
λ	4.498578	-3.900889	3.956493

偵測結果為第 188、242 及 322 筆資料為 IO 離群值，將其帶入模型配適，得到最終模型：

$$(1 - B^{12})(1 - B)Y_t = (1 + \hat{\theta}_1 B + \hat{\theta}_2 B^2)(1 + \hat{\Theta} B^{12})e_t$$

with IO. 188, IO. 242, IO. 322

	$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\Theta}$	IO.188	IO.242	IO.322
Estimator	-0.2193	-0.1523	-0.9065	2.4327	-2.0717	2.1066
Standard Error	0.0435	0.0433	0.0272	0.5521	0.5494	0.5499

由上表可知，所有參數估計值的正負兩倍標準差皆不包含 0，因此可判定第 188、242 及 322 筆資料為 IO 離群值。

iii. **SARIMA(1, 1, 1) × (0, 1, 1)₁₂**

$$(1 - B^{12})(1 - B)(1 - \hat{\phi}B)Y_t = (1 + \hat{\theta}B)(1 + \hat{\Theta}B^{12})e_t$$

	$\hat{\phi}$	$\hat{\theta}$	$\hat{\Theta}$
Estimator	0.4202	-0.6550	-0.9024
Standard Error	0.1087	0.0884	0.0286

由上表可知，所有參數估計值的正負兩倍標準差皆不包含 0，因此接著對該模型進行離群值檢測。

	IO[1]	IO[2]
n th	188	322
λ	4.471287	3.894541

偵測結果為第 188 及 322 筆資料為 IO 離群值，將其帶入模型配適，得到：

$$(1 - B^{12})(1 - B)(1 - \hat{\phi}B)Y_t = (1 + \hat{\theta}B)(1 + \hat{\Theta}B^{12})e_t$$

with IO. 188, IO. 322

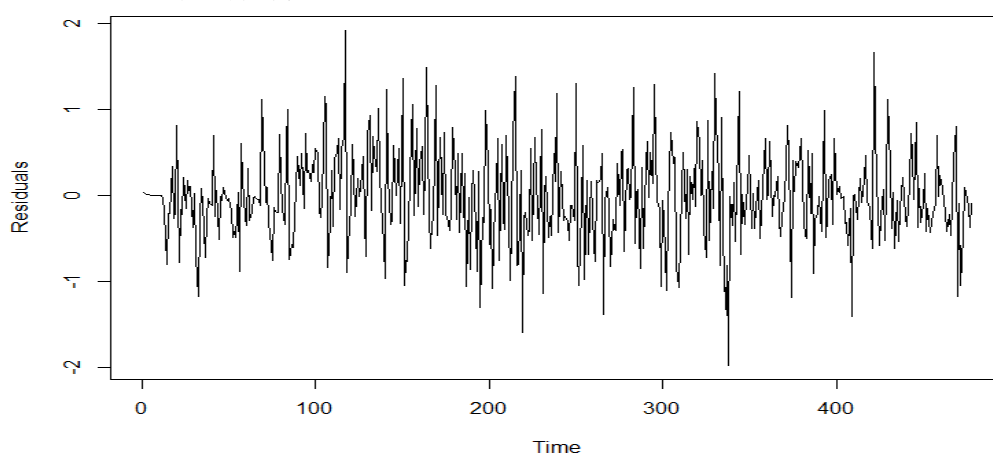
	$\hat{\phi}$	$\hat{\theta}$	$\hat{\Theta}$	IO.188	IO.322
Estimator	0.4385	-0.6700	-0.9113	2.4139	2.0776
Standard Error	0.1007	0.0805	0.0276	0.5601	0.5575

由上表可知，所有參數估計值的正負兩倍標準差皆不包含 0，因此可判定第 188 及 322 筆資料為 IO 離群值，為最終模型。

(6) 殘差模型診斷

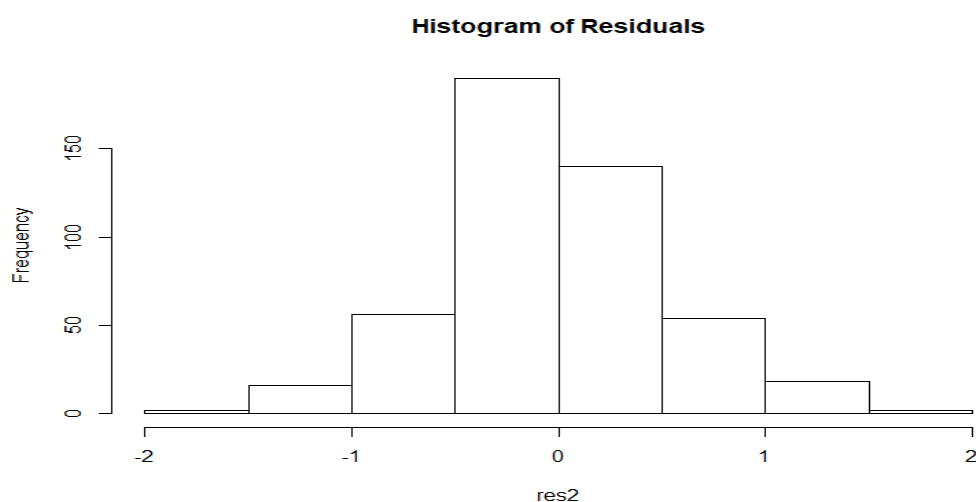
以 **SARIMA(0, 1, 2) × (0, 1, 1)₁₂**, with IO. 188, IO. 242, IO. 322 為例

i. 殘差得時間序列圖



殘差的時間序列圖可發現，殘差大致落在 0 的附近，無明顯的趨勢。

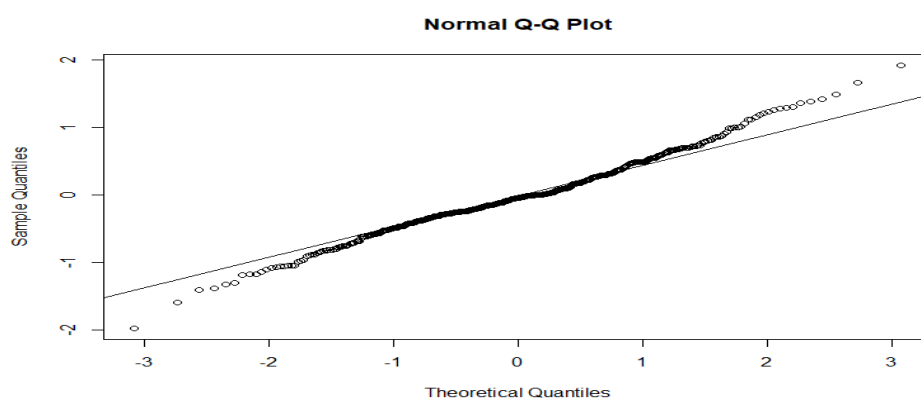
ii. 殘差平均是否為 0



One sample t-test	
t	-0.92474
df	477
p-value	0.3556
mean of x	-0.02283136
alternative hypothesis	Residual's mean is not equal to 0.

從殘差的直方圖可看出，殘差均勻地分布在零的左右；由 t-test 可得知 $p\text{-value} > 0.05$ ，不拒絕虛無假設，因此判斷殘差的平均為零，此檢定通過。

iii. 殘差平均是否為常態分布

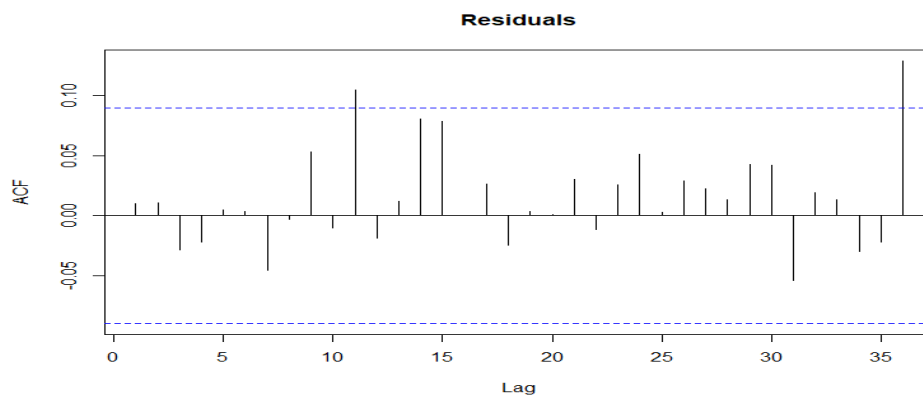


Shapiro-Wilk test	
w	0.98882
p-value	0.001036
alternative hypothesis	Residual is not a normal distribution.

Kolomogorov-Smirnov test	
D	0.066732
p-value	0.002807
alternative hypothesis	Residual is not a normal distribution.

透過 Q-Q Plot 可發現樣本百分位的兩端偏離常態線；由 Shapiro-Wilk test 與 Kolomogorov-Smirnov test 可得知二者的 $p\text{-value} < 0.05$ ，拒絕虛無假設，因此判斷殘差不為常態分布，此檢定不通過。

iv. 殘差的 ACF 是否為 0



Ljung-Box test		
χ^2	8.54	31.484
df	11	36
p-value	0.6643	0.6832
alternative hypothesis	At least one ACF is not equal to 0.	

由殘差的 ACF 圖可看出，在第 11 及 36 步時超出信賴界線，因此對此兩步數做 Ljung-Box test，可得知 $p\text{-value} > 0.05$ ，不拒絕虛無假設，因此判斷殘差的 ACF 為 0，此檢定通過。

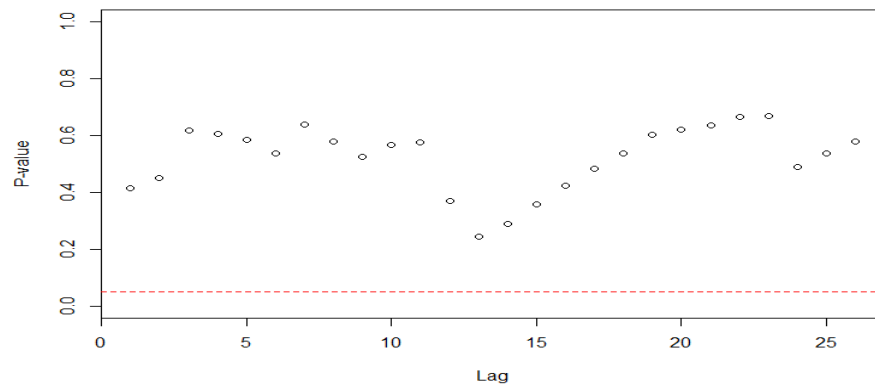
(7) 比較候選模型

將兩個候選模型的殘差檢定結果及 AIC 整理如下：

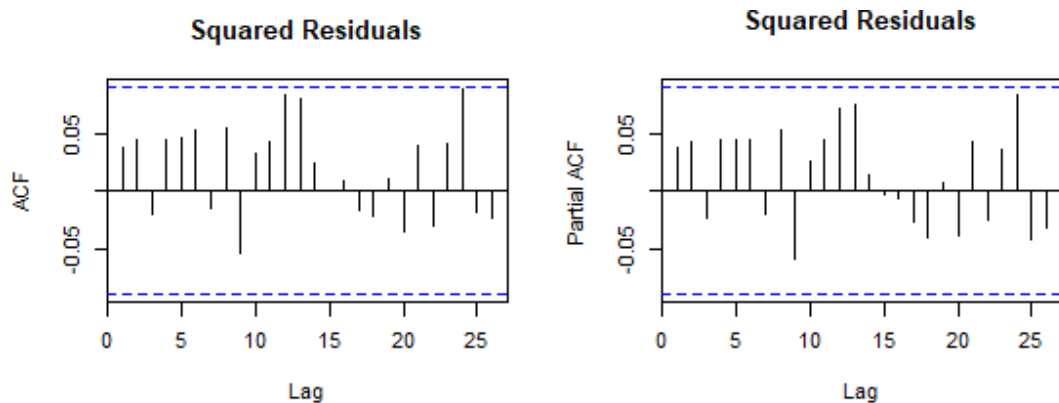
	SARIMA(0, 1, 2) × (0, 1, 1)₁₂ with IO. 188, IO. 242, IO. 322	SARIMA(1, 1, 1) × (0, 1, 1)₁₂ with IO. 188, IO. 322
殘差平均為 0	是	是
殘差為常態分布	否	否
殘差的 ACF 為 0	是	是
AIC	793.5547	805.5769

由於上述的兩個候選模型皆通過「殘差平均為 0」及「殘差的 ACF 為 0」，僅未通過「殘差為常態分布」的檢定，因此比較二者的 AIC，可看出 **SARIMA(0, 1, 2) × (0, 1, 1)₁₂**, with IO. 188, IO. 242, IO. 322 的 AIC 較小，判定其為較合適的模型，但常態檢定未通過，懷疑有 ARCH。

(8) 確認有無 ARCH 效應



由 McLeod Li Test 所有點皆落在信賴線外，判斷此資料沒有 ARCH 效應，無需建立 GARCH 模型。



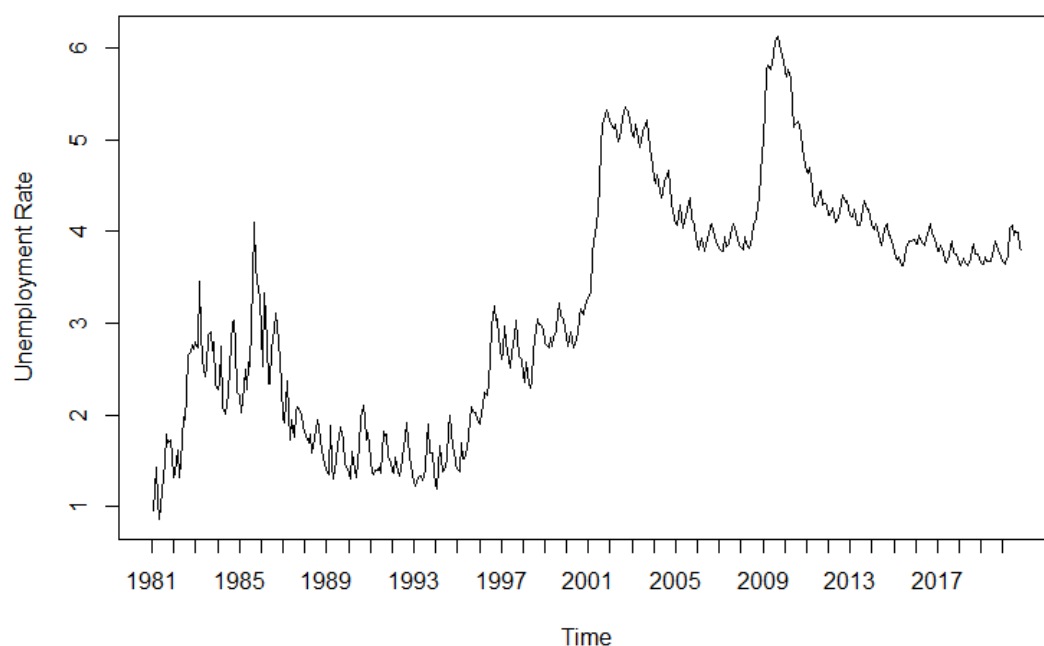
從殘差平方的 ACF 圖（左上）及 PACF 圖（右上）可看出，所有步數皆無超出信賴界線，再次確認無需建立 GARCH 模型。

(9) 小結

最終的模型沒有 Garch 效應，也沒有通過常態檢定，因此判斷消費者物價指數的資料本身並非為常態分布。

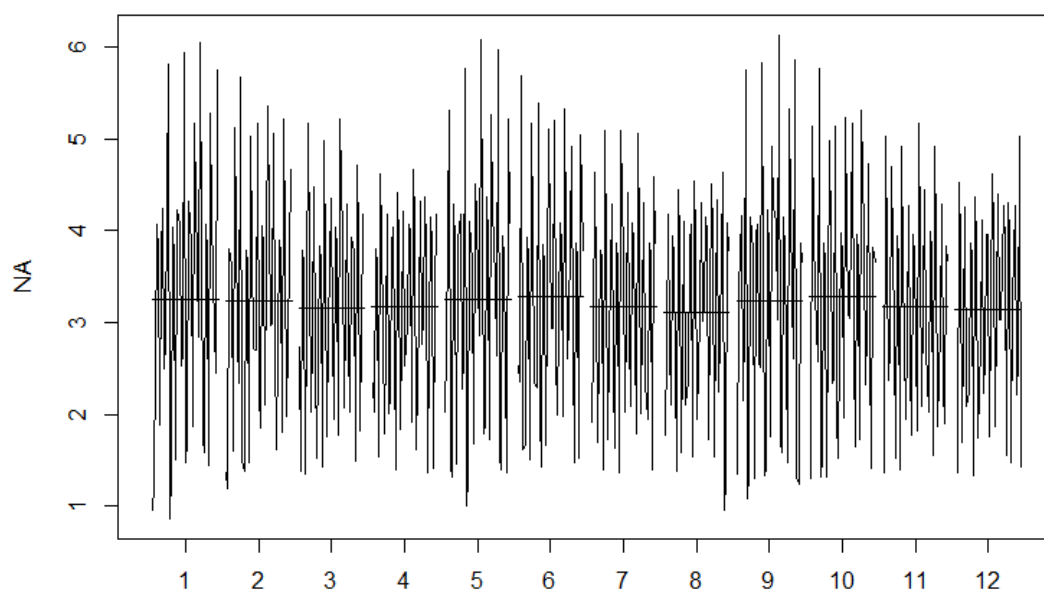
2. 失業率的模型建置

(1) 繪製失業率的時間序列圖



由上圖可看出，整體資料有呈現上升的趨勢，且 1986 至 2002 年、2002 至 2010 年皆有逐年下降再大幅度上升的趨勢。

(2) 繪製失業率的 Month Plot



由上圖可看出，各月份的分布差異並不顯著，因此尚無法判斷此資料是否具有季節效應，需待後續繪製 ACF 及 PACF 圖以利判斷。

(3) 確認資料的平穩

為了確認資料是否需要差分，因此做 Dickey-Fuller Test 及 KPSS Test，以檢定的方式來判斷。

i. Dickey-Fuller Test

H_0 : Non-Stationary. (Need difference.) H_1 : Stationary. (Do not need difference.)	
Dickey-Fuller	-2.198
p-value	0.4943
Conclusion	Do not reject H_0 , need difference.

ii. KPSS Test

H_0 : Stationary. (Do not need difference.) H_1 : Non-Stationary. (Need difference.)	
KPSS Level	5.0493
p-value	0.01
Conclusion	Reject H_0 , need difference.

從兩個檢定方法都得出資料不平穩，需差分的結論，因此再利用上述兩個檢定判斷是否需要二階差分。

i. Dickey-Fuller Test

H_0 : Non-Stationary. (Need difference.) H_1 : Stationary. (Do not need difference.)	
Dickey-Fuller	-8.7133
p-value	0.01
Conclusion	Reject H_0 , don't need difference.

ii. KPSS Test

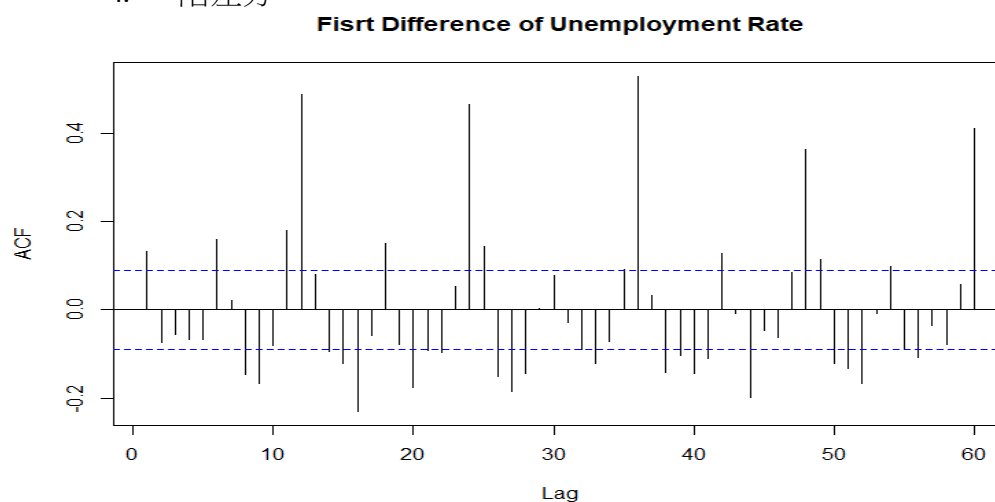
H_0 : Stationary. (Do not need difference.) H_1 : Non-Stationary. (Need difference.)	
KPSS Level	0.093153
p-value	0.1
Conclusion	Do not reject H_0 , don't need difference.

由上述的結果可知此資料一階差分後即平穩，無需做二階差分。

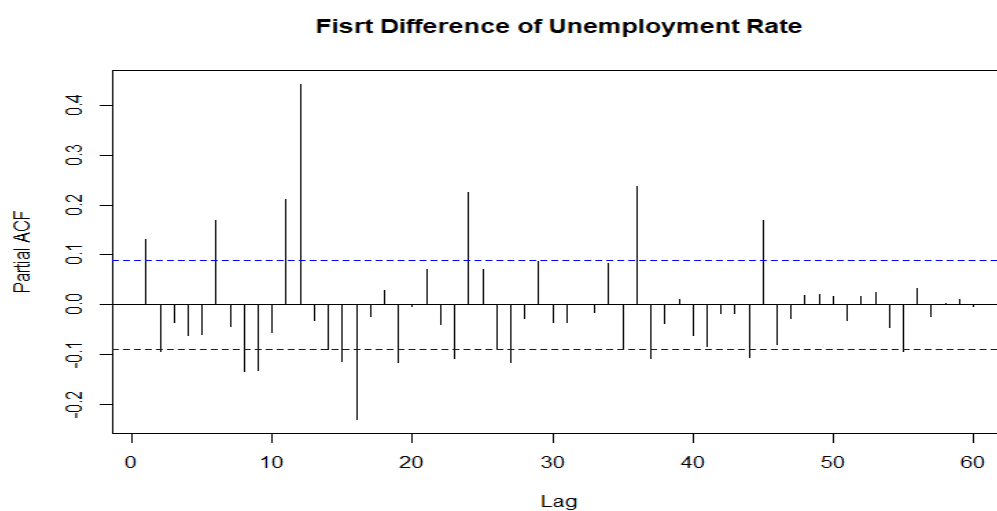
(4) 選擇候選模型

透過觀察資料的 ACF、PACF 及 EACF 圖，選擇合適的候選模型。

i. 一階差分



從 ACF 圖（上圖）可看出，第 9 步時仍有顯著性，且在第 12、24、36、48 及 60 步時也皆超出信賴界線。

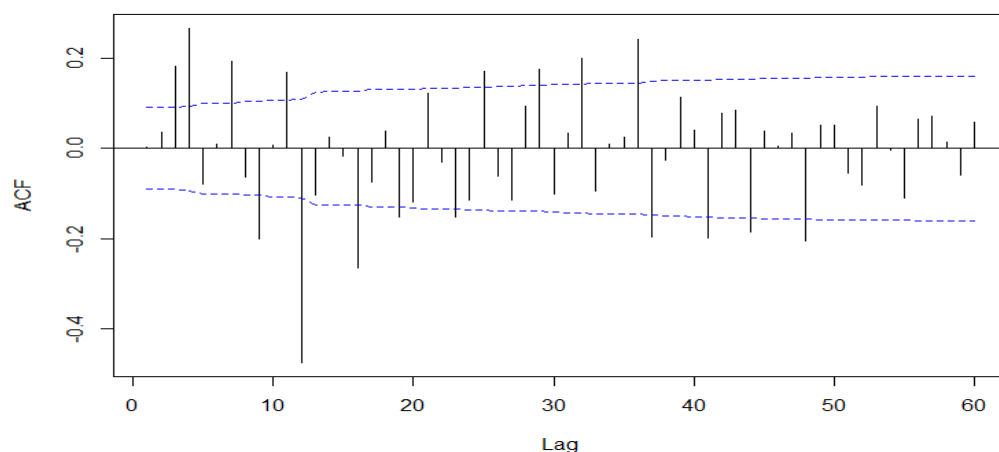


從 PACF 圖（上圖）可看出，第 9 步時仍有顯著性，且在第 12、24、36、48 及 60 步時也皆超出信賴界線。

綜合上述的 ACF 及 PACF 圖可看出，該資料應具有季節效應，因此後續將對此資料進行一階差分及季節性一階差分。

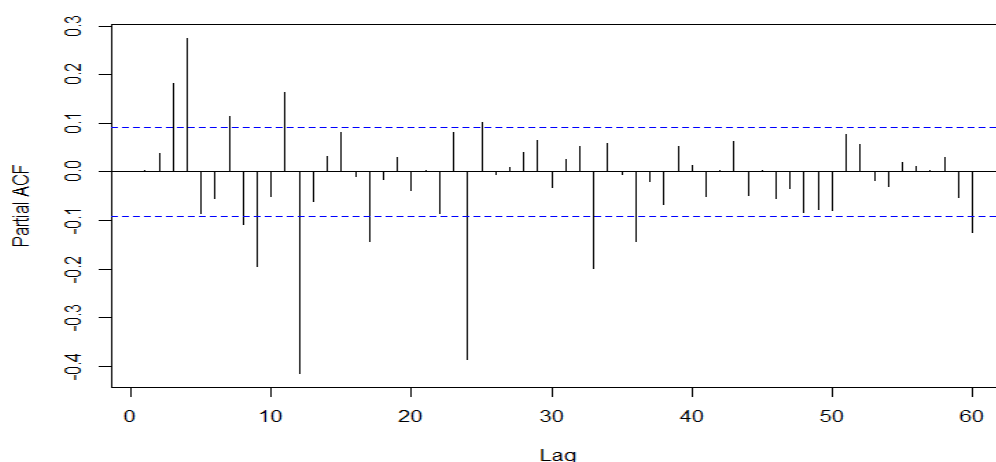
ii. 一階差分與季節性一階差分

First Difference and Seasonal First Difference of Unemployment Rate



從 ACF 圖（上圖）可看出，在第 4 步時截斷，而在第 11、12 步後仍有些許步數超出信賴界線，但由於影響資料的步數過大，因此僅挑選 $SARIMA(0,1,4) \times (0,1,1)_{12}$ 候選模型。

First Difference and Seasonal First Difference of Unemployment Rate



從 PACF 圖（上圖）可看出，第 9 步時仍有顯著性，且在第 11、12、24、25、33、36 及 60 步時也皆超出信賴界線，由於影響資料的步數過大，因此暫不挑選候選模型。

AR/MA															
		0	1	2	3	4	5	6	7	8	9	10	11	12	13
0		o	o	x	x	o	o	x	o	x	o	x	x	o	
1		o	o	o	x	o	o	x	o	x	o	o	x	x	o
2		x	o	o	x	o	o	o	o	o	o	o	x	x	o
3		x	x	o	x	o	o	o	o	o	o	o	x	o	o
4		x	x	x	x	o	o	o	o	o	o	o	x	o	o
5		x	x	o	x	o	o	o	o	o	o	o	x	x	o
6		x	x	o	x	o	x	o	o	o	o	o	x	x	o
7		x	x	x	x	x	x	x	o	o	o	o	x	x	o

從 EACF 圖（上圖）可得知候選模型為 $SARIMA(0,1,0) \times (0,1,1)_{12}$ ，因為 $ARMA(0,0)$ 的三角區域涵蓋較多「o」。

綜合上述，共得 2 個候選模型，以下將進行模型的參數估計。

(5) 估計候選模型參數

i. **SARIMA(0, 1, 4) × (0, 1, 1)₁₂**

$$(1 - B^{12})(1 - B)Y_t = (1 + \hat{\theta}_1 B + \hat{\theta}_2 B^2 + \hat{\theta}_3 B^3 + \hat{\theta}_4 B^4)(1 + \hat{\Theta} B^{12})e_t$$

	$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\theta}_3$	$\hat{\theta}_4$	$\hat{\Theta}$
Estimator	0.0483	0.0341	0.0792	0.1884	-0.8135
Standard Error	0.0462	0.0460	0.0425	0.0474	0.0271

由上表可知， $\hat{\theta}_1$ 、 $\hat{\theta}_2$ 及 $\hat{\theta}_3$ 估計值的正負兩倍標準差皆包含 0，但未避免資料須進行離群值的參數估計而影響此三個參數估計，因此先接著對該模型進行離群值檢測。

	AO[1]	AO[2]	AO[3]	AO[4]
n th	51	68	80	86
λ	5.121192	-4.484407	-4.675084	-5.319893

	IO[1]	IO[2]	IO[3]	IO[4]	IO[5]	IO[6]	IO[7]
n th	29	51	55	57	62	86	336
λ	-4.125	7.912	4.331	-4.175	4.191	-5.371	3.977

上述兩個表格為偵測結果，共有 4 筆資料為 AO 離群值及 7 筆資料為 IO 離群值，將其帶入模型配適並剔除不具顯著性得參數估計，可得到最終模型：

$$(1 - B^{12})(1 - B)Y_t = (1 + \hat{\theta}_1 B + \hat{\theta}_2 B^2 + \hat{\theta}_3 B^3 + \hat{\theta}_4 B^4)(1 + \hat{\Theta} B^{12})e_t$$

with Apr. 2020, IO. 29, IO. 51, IO. 55, IO. 57, IO. 62, IO. 86, IO. 87, IO. 336

	$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\theta}_3$	$\hat{\theta}_4$	$\hat{\Theta}$	Apr.2020	IO.29
Estimator	0.1102	0.0824	0.0812	0.1777	-0.79	0.1567	-0.4395
Standard Error	0.0427	0.0415	0.0363	0.0415	0.0243	0.0750	0.1253
	IO.51	IO.55	IO.57	IO.62	IO.86	IO.87	IO.336
Estimator	0.9096	0.5234	-0.510	-.4911	-0.573	0.4778	0.3977
Standard Error	0.1176	0.1248	0.1191	0.1160	0.1148	0.1177	0.1144

ii. **SARIMA(0, 1, 0) × (0, 1, 1)₁₂**

$$(1 - B^{12})(1 - B)Y_t = (1 + \hat{\Theta}B^{12})e_t$$

	$\hat{\Theta}$
Estimator	-0.8022
Standard Error	0.0227

由上表可知，此參數估計值的正負兩倍標準差不包含 0，因此接著對該模型進行離群值檢測。

	AO[1]
n th	80
λ	-4.128

	IO[1]	IO[2]	IO[3]	IO[4]	IO[5]
n th	29	51	55	57	62
λ	-4.067	6.965	5.491	-4.175	4.021
	IO[6]	IO[7]	IO[8]	IO[9]	IO[10]
n th	86	87	92	122	336
λ	-4.776	4.375	-3.927	-3.913	4.231

上述兩個表格為偵測結果，共有 1 筆資料為 AO 離群值及 10 筆資料為 IO 離群值，將其帶入模型配適並剔除不具顯著性得參數估計，可得到最終模型：

$$(1 - B^{12})(1 - B)Y_t = (1 + \hat{\Theta}B^{12})e_t$$

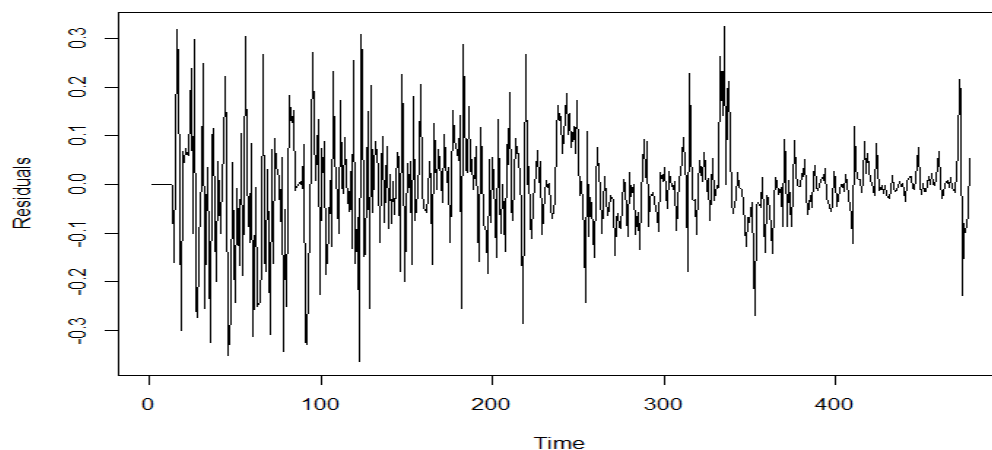
with Aug. 1987, IO. 29, IO. 51, IO. 55, IO. 57, IO. 62, IO. 86, IO. 87, IO. 92, IO. 122, IO. 336

	$\hat{\Theta}$	Aug.1987	IO.29	IO.51	IO.55	IO.57
Estimator	-0.8022	-0.1858	-0.4784	0.8153	0.6632	-0.4113
Standard Error	0.0227	0.0784	0.1250	0.1189	0.1187	0.1186
	IO.62	IO.86	IO.87	IO.92	IO.122	IO.336
Estimator	0.4217	-0.5294	0.445	-0.3967	-0.3897	0.4673
Standard Error	0.1173	0.1160	0.116	0.1170	0.1152	0.1149

(6) 殘差模型診斷

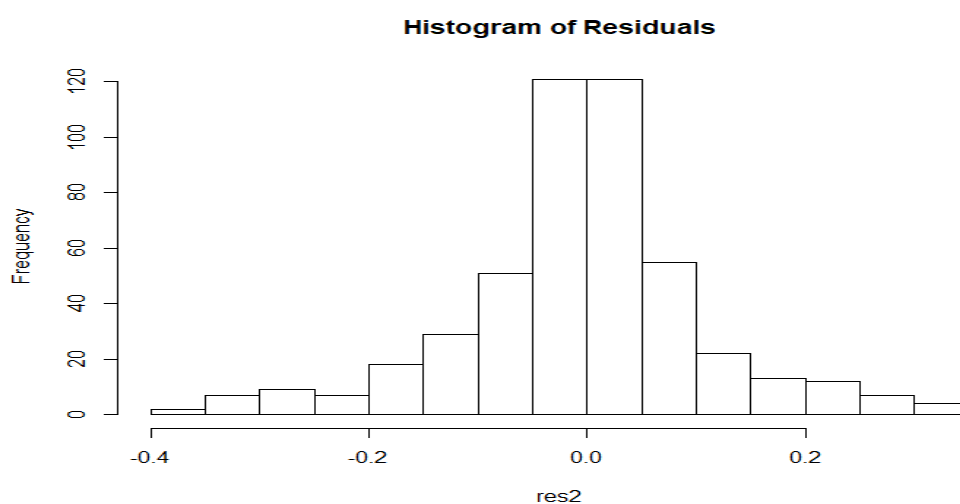
以 **SARIMA(0, 1, 4) × (0, 1, 1)₁₂** with Apr. 2020, IO. 29, IO. 51, IO. 55, IO. 57, IO. 62, IO. 86, IO. 87, IO. 336 為例

i. 殘差得時間序列圖



殘差的時間序列圖可發現，殘差大致落在 0 的附近，無明顯的趨勢。

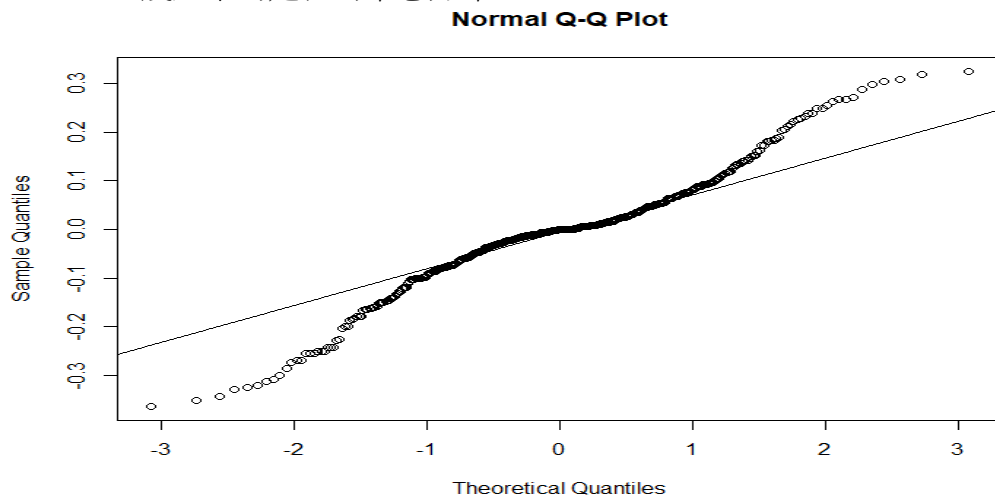
ii. 殘差平均是否為 0



One sample t-test	
t	-1.142
df	477
p-value	0.254
mean of x	-0.005816742
alternative hypothesis	Residual's mean is not equal to 0.

從殘差的直方圖可看出，殘差均勻地分布在零的左右；由 t-test 可得知 $p\text{-value} > 0.05$ ，不拒絕虛無假設，因此判斷殘差的平均為零，此檢定通過。

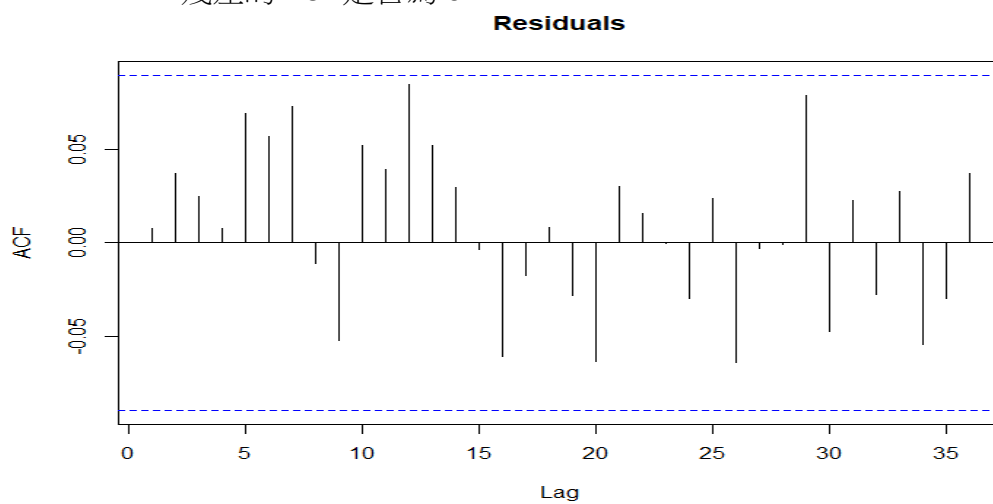
iii. 殘差平均是否為常態分布



Shapiro-Wilk test	
w	0.95771
p-value	1.778×10^{-10}
alternative hypothesis	Residual is not a normal distribution.

透過 Q-Q Plot 可發現樣本百分位的兩端偏離常態線；由 Shapiro-Wilk test 可得知 $p\text{-value} < 0.05$ ，拒絕虛無假設，因此判斷殘差不為常態分布，此檢定不通過。

iv. 殘差的 ACF 是否為 0



由殘差的 ACF 圖可看出，所有步數皆未超出信賴界線，因此不需做 Ljung-Box test，即可判斷殘差的 ACF 為 0，此檢定通過。

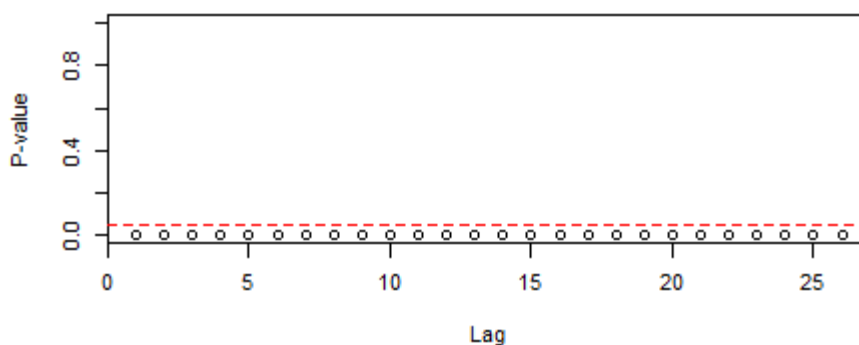
(7) 比較候選模型

將兩個候選模型的殘差檢定結果及 AIC 整理如下：

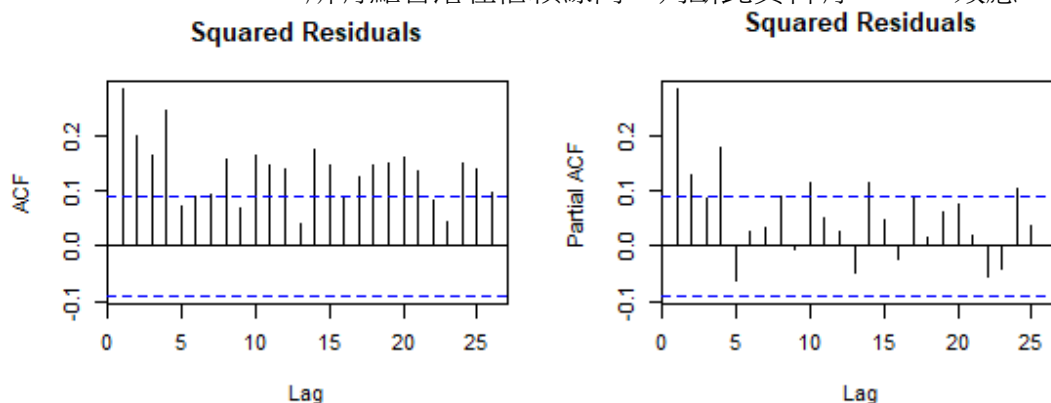
	SARIMA(0, 1, 4) × (0, 1, 1)₁₂ with Apr. 2020, IO. 29, IO. 51, IO. 55, IO. 57, IO. 62, IO. 86, IO. 87, IO. 336	SARIMA(0, 1, 0) × (0, 1, 1)₁₂ with Aug. 1987, IO. 29, IO. 51, IO. 55, IO. 57, IO. 62, IO. 86, IO. 87, IO. 92, IO. 122, IO. 336
殘差平均為 0	是	是
殘差為常態分布	否	否
殘差的 ACF 為 0	是	否
AIC	-668.75	-656.32

由於上述的兩個候選模型皆通過「殘差平均為 0」及未通過「殘差為常態分布」的檢定，而「殘差的 ACF 為 0」之檢定僅有模型 **SARIMA(0, 1, 4) × (0, 1, 1)₁₂**, with Apr. 2020, IO. 29, IO. 51, IO. 55, IO. 57, IO. 62, IO. 86, IO. 87, IO. 336 通過，因此判定為較合適的模型，但未通過常態檢定，懷疑有 ARCH 效應。

(8) 確認有無 ARCH 效應

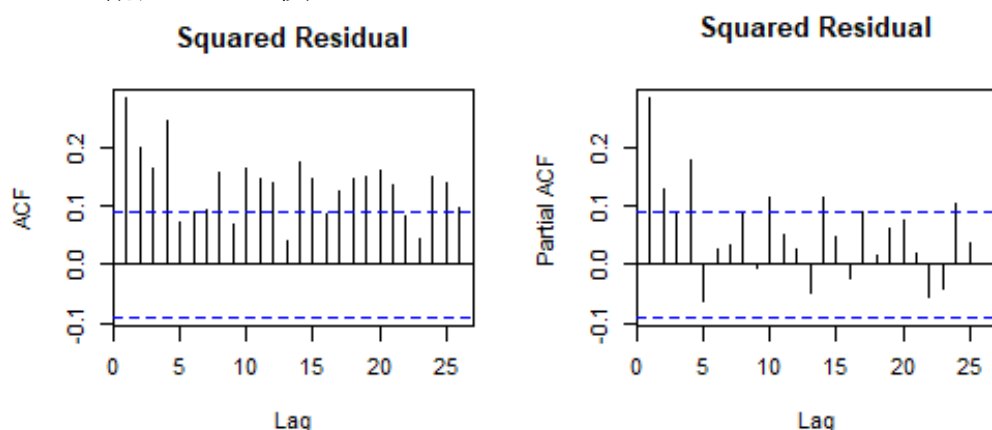


McLeod Li Test 所有點皆落在信賴線內，判斷此資料有 ARCH 效應。



從殘差平方的 ACF 圖（左上）可看出大部分的步數都超出信賴界，而 PACF 圖（右上）可得知，1、2、4 步超出信賴界，再次確認有 ARCH 效應，因此殘差需建立 GARCH 模型。

i. 配適 GARCH 模型



殘差平方的 ACF 圖（左上）可看出大部分的步數都超出信賴界，無法選出適當的候選模型，而 PACF 圖（右上）可得知，1、2、4 步超出信賴界，因此候選模型可選至 GARCH（4，0）。

	0	1	2	3	4	5	6	7	8	9	10	11	12	13
0	x	x	x	x	o	o	o	x	o	x	x	x	o	x
1	x	o	o	x	x	o	o	x	o	o	o	x	o	x
2	x	o	o	x	x	o	o	o	o	o	o	o	x	x
3	x	x	x	x	o	o	o	o	o	o	o	o	o	o
4	x	o	x	x	x	x	o	o	o	o	o	o	o	o
5	x	o	o	x	x	o	o	o	o	o	o	o	o	o
6	x	o	o	x	o	x	o	o	o	o	o	o	o	o
7	x	x	x	x	o	x	o	o	o	o	o	o	o	o

從殘差平方的 EACF 圖可發現可選用的候選模型步數較大，因此暫時不當作候選模型的參考。

ii. GARCH 模型參數估計：GARCH（4，0）。

在配適 GARCH 模型時，有考慮 GARCH（1，0）和 GARCH（2，0），但兩者皆在「殘差的 ACF 為 0」之檢定未通過，因此不採納。

以下為 SARIMA(0,1,4) × (0,1,1)₁₂, with Apr. 2020, IO. 29, IO. 51, IO. 55, IO. 57, IO. 62, IO. 86, IO. 87, IO. 336 模型加上 GARCH(4,0) 模型作為失業率的最終模型，模型如下：

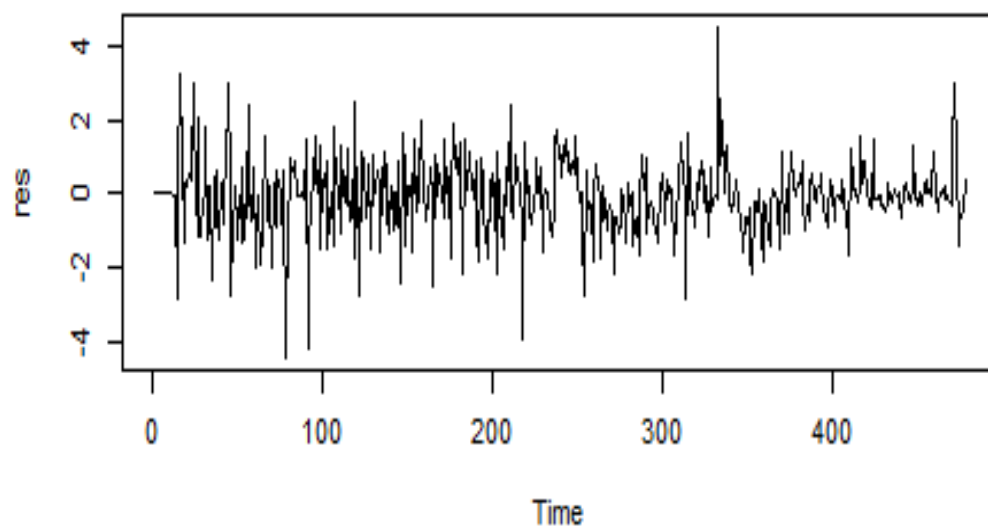
$$\begin{cases} (1 - B^{12})(1 - B)\gamma_t = (1 + \hat{\theta}_1 B + \hat{\theta}_2 B^2 + \hat{\theta}_3 B^3 + \hat{\theta}_4 B^4)(1 + \hat{\theta} B^{12})e_t, \\ \text{with Apr. 2020, IO. 29, IO. 51, IO. 55, IO. 57, IO. 62, IO. 86, IO. 87, IO. 336} \\ \gamma_t = \sigma_t e_t, \quad e_t \sim WN(0,1) \\ \sigma_t^2 = \hat{\omega} + \hat{\alpha}_1 \gamma_{t-1}^2 + \hat{\alpha}_2 \gamma_{t-2}^2 + \hat{\alpha}_3 \gamma_{t-3}^2 + \hat{\alpha}_4 \gamma_{t-4}^2 \end{cases}$$

	$\hat{\omega}$	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\alpha}_3$	$\hat{\alpha}_4$
Estimator	0.0032	0.4288	0.2172	0.0505	0.1806
P-value	$< 4.89 \times 10^{-9}$	$< 5.14 \times 10^{-5}$	0.0010	0.342	0.004

※ SARIMA(0,1,4) × (0,1,1)₁₂ 模型之參數估計同第 18 頁。

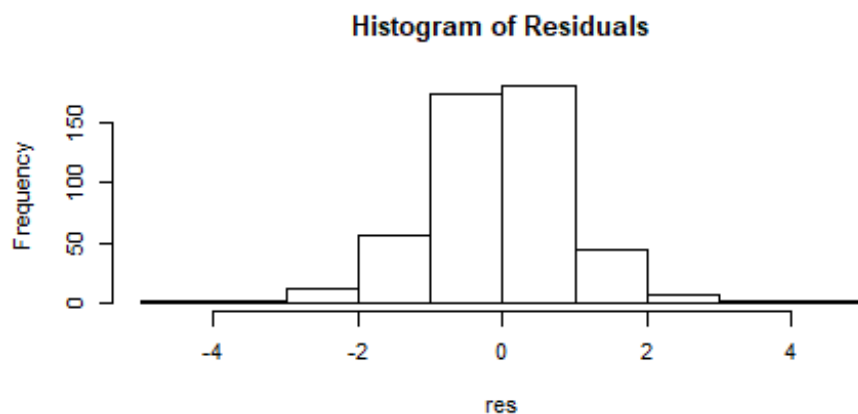
iii. GARCH (4, 0) 模型診斷

iii.1 殘差得時間序列圖



殘差的時間序列圖可發現，殘差大致落在 0 的附近，無明顯的趨勢。

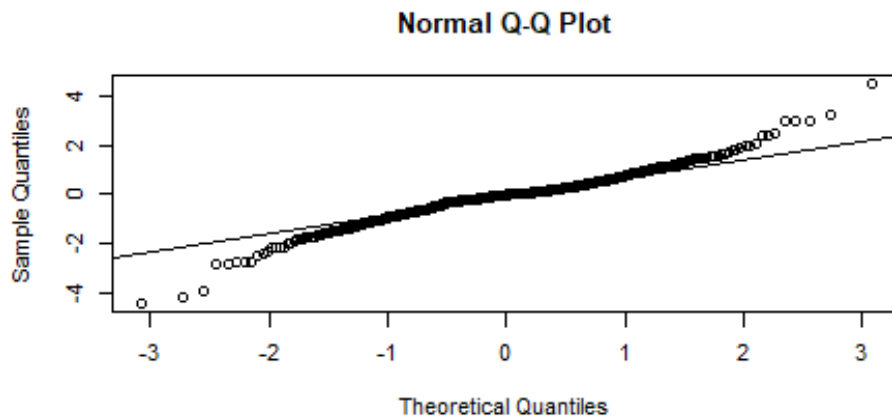
iii.2 殘差平均是否為 0



One sample t-test	
t	-1.6241
df	477
p-value	0.105
mean of x	-0.07415617
alternative hypothesis	Residual's mean is not equal to 0.

從殘差的直方圖可看出，殘差均勻地分布在零的左右；由 t-test 可得知 $p\text{-value} > 0.05$ ，不拒絕虛無假設，因此判斷殘差的平均為零，此檢定通過。

iii.3 殘差平均是否為常態分布

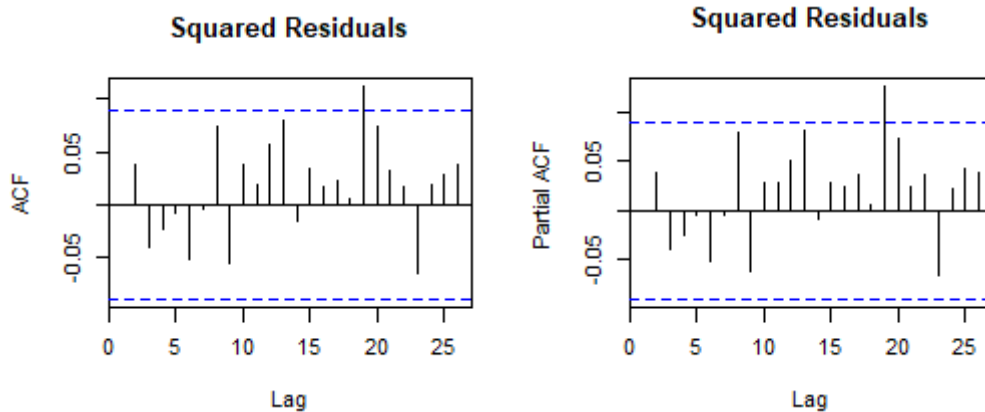


Shapiro-Wilk test	
w	0.96187
p-value	8.519e-10
alternative hypothesis	Residual is not a normal distribution.

Kolomogorov-Smirnov test	
D	0.094578
p-value	0.003866
alternative hypothesis	Residual is not a normal distribution.

透過 Q-Q Plot 可發現樣本百分位的兩端偏離常態線；由 Shapiro-Wilk test 和 Kolomogorov-Smirnov test 可得知 $p\text{-value} < 0.05$ ，拒絕虛無假設，因此判斷殘差不為常態分布，此檢定不通過。

iii.4 殘差與殘差平方的 ACF 是否為 0



由殘差的 ACF 圖（左上）可看出，步數 19 超出信賴界線，而殘差平方的 ACF 圖（右上）可看，步數 19 超出信賴界線，因此需做 Ljung-Box Test。

iii.4 殘差與殘差平方的 ACF 是否為 0 (續 Ljung-Box Test)

Ljung-Box test				
χ^2	4.8549	17.424	28.771	33.777
df	5	12	20	26
p-value	0.4338	0.1343	0.09237	0.1407
alternative hypothesis	At least one ACF is not equal to 0.			
data	Residuals			

Ljung-Box test	
χ^2	20.174
df	19
p-value	0.3842
alternative hypothesis	At least one ACF is not equal to 0.
data	Squared Residuals

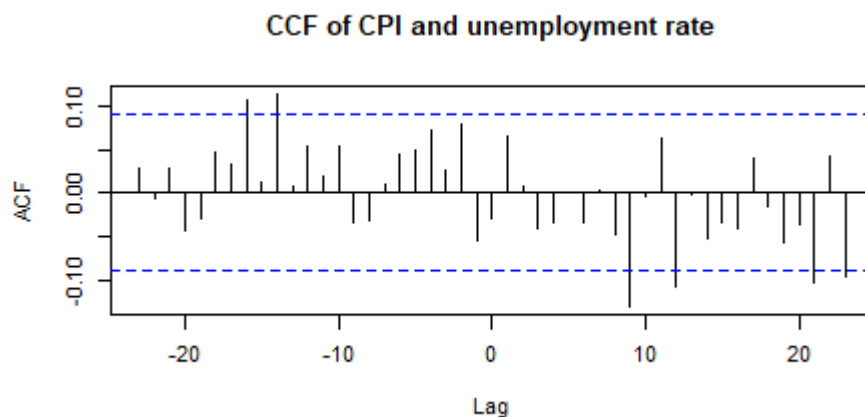
由 Ljung-Box test 可得知所有的 $p\text{-value} > 0.05$ ，不拒絕虛無假設，因此判斷殘差的 ACF 為 0，此檢定通過。

(9) 小結

最終的模型即使配適完成 Garch 模型，也沒有通過常態檢定，因此判斷失業率的資料本身並非為常態分布。

五、相關性分析

為討論消費者物價指數與失業率間的相關性，首先繪製 CCF 圖。



從 CCF 圖觀察有無延遲步數的相關性，可以發現步數為 -16、-14、9、12 時都超出信賴界線。步數為-16 或-14 時，代表消費者物價指數延遲-16 或-14 步影響失業率；步數為 9 或 12 時，代表失業率延遲 9 或 12 步影響消費者物價指數。為以上四種可能配適迴歸模型。

(1) Lag = 9

$$\varepsilon_t = \beta_0 + \beta_1 e_{t-9} + \eta_t, \quad e_{t-9} : \text{失業率}, \varepsilon_t : \text{消費者物價指數}$$

	$\hat{\beta}_0$	$\hat{\beta}_1$
估計值	-0.02555	-0.02643
P-value	0.311	0.298

迴歸模型的參數估計從 p-value 可以都大於 0.05，不拒絕虛無假設，因此判定沒有顯著性。

(2) Lag = 12

$$\varepsilon_t = \beta_0 + \beta_1 e_{t-12} + \eta_t, \quad e_{t-12} : \text{失業率}, \varepsilon_t : \text{消費者物價指數}$$

	$\hat{\beta}_0$	$\hat{\beta}_1$
估計值	-0.02070	-0.03563
P-value	0.414	0.162

迴歸模型的參數估計從 p-value 可以都大於 0.05，不拒絕虛無假設，因此判定沒有顯著性。

(3) lag = -14

$\varepsilon_t = \beta_0 + \beta_1 e_{t-14} + \eta_t$ ， e_{t-14} ：消費者物價指數， ε_t ：失業率

	$\hat{\beta}_0$	$\hat{\beta}_1$
估計值	-0.07598	-0.01283
P-value	0.107	0.883

迴歸模型的參數估計從 p-value 可以都大於 0.05，不拒絕虛無假設，因此判定沒有顯著性。

(4) lag = -16

$\varepsilon_t = \beta_0 + \beta_1 e_{t-16} + \eta_t$ ， e_{t-16} ：消費者物價指數， ε_t ：失業率

	$\hat{\beta}_0$	$\hat{\beta}_1$
估計值	-0.07103	0.02203
P-value	0.130	0.799

迴歸模型的參數估計從 p-value 可以都大於 0.05，不拒絕虛無假設，因此判定沒有顯著性。

綜合上述，所有模型皆都沒有顯著性，所以推斷消費者物價指數與失業率之間無相關性。

六、結論

由以上的資料建模與相關性分析可得知消費者物價指數與失業率之間無相關性，可能原因為消費者物價指數是以消費者的立場衡量財貨及勞務的價格，是反映通貨膨脹的重要指標之一，但不能延伸推論市場上因為物價的上升或下降影響失業率。

七、參考資料

1. 消費者物價指數資料來源：中華民國統計資料網
<https://www.stat.gov.tw/np.asp?ctNode=485&mp=4>
2. 失業率資料來源：中華民國統計資料網
<https://www.stat.gov.tw/np.asp?ctNode=514&mp=4>
3. 消費者物價指數維基百科

八、課堂心得與建議

1. 心得

修了一學期的時間序列分析課程，讓我們更加了解時間效應對於資料分析的重要性，使我們從僅會配適一般迴歸模型來分析資料的相關性，到期末時已能利用時間序列分析的方法，將資料配適合適的 **ARIMA**、**SARIMA** 及 **GARCH** 模型以進行更進一步的相關性分析與預測，我們覺得受益良多且有助於爾後我們進行財務相關的時間序列資料之分析，在此謝謝老師的辛勞及教導。

2. 建議

在期中前的課程安排，我們覺得大致上沒有問題，但期中過後的課程本應著重於實證資料的分析及應用，但卻花了大多的時間在理論推導上，不但對於期末報告無明顯益處，大部分同學應該也都無法有效吸收，因此我們認為後半段的課程老師應需調整教學的重心。在進行實證資料分析時，我們認為可以給予同學一張完整的分析時間序列資料的步驟表，以免同學在學習的過程中會不知道自己正在做什麼，此表應能有助於同學更加了解各步驟間的銜接與關係，在製作期末報告時，也能一一檢視是否有任何疏漏以減少錯誤率。

關於末段課程(CH10~12)的建議，我們認為應著重於 **R** 程式及實證資料的分析及應用，大部分的理論推導對於期末報告並沒有太大的幫助，倘若末段課程能調整為：先簡單帶過理論→再著重於 **R** 及實證資料，若之後還有充足的時間或大部份學生對於理論的推導有興趣，最後再進行更進一步的理論推導。