

# 逢甲大學學生報告 ePaper

來台旅遊人數與國人出國人數之時間序列預測分析

Predictive Analysis of Time Series Modeling for the Number of  
Tourists to Taiwan and the Number of Outbound Travelers

作者：陳志豪

系級：統計三乙

學號：D0781161

開課老師：劉峰旗

課程名稱：預測分析

開課系所：統計學系

開課學年：109 學年度 第 2 學期

## 中文摘要

隨著時代的變遷和近年來國民所得的增加，以及生活水準的提升，國內外重視休閒的人口逐漸增加，再加上政府目前實施週休二日，許多家庭外出旅遊的比例越來越高，出國觀光旅遊的人數更是快速上升，而來台的旅客人數也是只增不減，因此造就了台灣觀光產業的蓬勃發展，為國內市場帶來不少的經濟收益。為了解來台旅客人數的成長趨勢，以及國人出國旅遊狀況，本文運用在「預測分析」課程所學習到的方法進行時間序列模型的探討，並且預測未來一年來台遊客人數，以作為我國觀光產業發展的參考。

本文採用交通部觀光局觀光統計資料庫所提供的來台旅遊人數及國人出國人數進行分析，資料範圍自 2010 年 1 月至 2018 年 12 月，共計 108 筆月資料，並保留最後 12 筆進行預測。本文採用時間序列迴歸法、指數平滑法、ARIMA 方法來進行模型配適，並透過殘差分析找出最佳模型。最後，再以 MAE、MSE 及 MAPE 準則來評估各模型的預測效果，並選出最佳模型。本文分析結果顯示，ARIMA 方法在三個準則之下皆有最佳的表現，該模型可做為來台旅客人數與國人出國人數之預測模型。

**關鍵字：**時間序列迴歸、指數平滑法、ARIMA 模型、來台旅客人數、國人出國人數

## Abstract

With the changes of the times, the increase in national income in recent years, and the improvement of living standards, the number of people paying attention to leisure at home and abroad has gradually increased. In addition, the government currently implements a two-day weekly holiday, and the proportion of many families going out for tourism is increasing. The number of people going abroad for sightseeing has increased directly, and the number of tourists coming to Taiwan has only increased. Therefore, this phenomenon has caused the flourishing development of Taiwan's tourism industry. The tourism industry will be the main force of Taiwan's future economic development. Visitors to Taiwan and the market. So I want to use the forecasting methods learned in this lesson of forecast analysis to find the best time series model, and predict the number of tourists in Taiwan in the next year, and find out how to increase the number of tourists in Taiwan, and hope this This report can be used as a reference for the Tourism Bureau.

The source of the data is the tourism statistics database of the Tourism Bureau of the Ministry of Transport. The analysis time is from January 2010 to December 2018, and the last 12 records are retained for forecasting. At the beginning of the report, confirm whether the time series graph shows a trend, or whether the variance and the average are stable. Then use time series regression, exponential smoothing, and ARIMA analysis to fit the model. Finally, use the three criteria of MAE, MSE and MAPE to evaluate which of these three methods is the best and select the best model. The analysis results show that the fit model of the ARIMA method is the best model.

**Keyword :** Time series regression, exponential smoothing method, ARIMA model, number of tourists to Taiwan, number of outbound travelers.

## 目錄

第一章 諸論.....	4
第一節 研究背景.....	4
第二節 研究動機.....	8
第三節 研究目的.....	9
第四節 文獻探討及回顧.....	9
第二章 研究方法.....	12
第一節 資料敘述和使用方法說明.....	12
第二節 研究流程.....	13
第三章 研究模型探討.....	14
第一節 時間序列回歸法(Time series regression).....	14
第二節 指數平滑法(Exponential smoothing).....	19
第三節 ARIMA 分析法.....	24
第四節 最佳模型.....	32
第四章 結論與建議.....	33
參考文獻.....	34

## 第一章 諸論

### 第一節 研究背景

2020 年正值 COVID-19 (2019 新冠狀肺炎病毒) 疫情重創全球，我國觀光產業也遭受影響，不僅民眾因恐懼不敢到國外旅遊，連國內旅遊業也大受影響，而且台灣消費及生活產業出現了較過往數年更為劇烈的動盪變化，尤其是旅遊休憩業受到相當大衝擊，受影響的產業橫跨交通運輸、飯店、餐飲、免稅店、旅行社等業者。根據交通部觀光局 109 年台灣旅遊狀況調查報告，報告中表示 109 年整年 12 歲及以上國人的國內旅遊總次數為 1.4 億次，較 108 年負成長 15.54%，旅遊支出也減少近新台幣 450 億元，以上種種跡象顯示我國觀光業深受新冠肺炎疫情的影響。

首先，針對交通部觀光局統計資料顯示，由下表 1 可以看出在 2010 至 2018 這 9 年間來台旅客的平均人數為 9169319 人，代表每年會有平均 9169319 人來台旅遊觀光，標準差為 2237498 人，而國人出國的平均人數則為 12930875 人，代表每年平均會有 12930875 人出國旅遊，標準差為 2923628 人。由平均數來觀察，每年國人出國的平均人數是大於來台旅客人數的，由標準差來看，國人出國人數的標準差也是大於來台旅客人數的標準差，所以代表國人出國旅遊的波動程度相較於來台的旅客，是比較大的。

表 1 2010 年至 2018 年來台旅客人數、國人出國人數、外匯收入敘述性統計

	最大值	最小值	平均數	中位數	標準差
來台旅客人數	11864105	5567277	9169319	10174994	2237498
國人出國人數	17101335	9415074	12930875	12513806	2923628
旅館住用率(%)	54.14	42.91	49.29	49.1	3.566254
外匯收入(美元)	14620000000	8719000000	12670000000	12850000000	1843239844

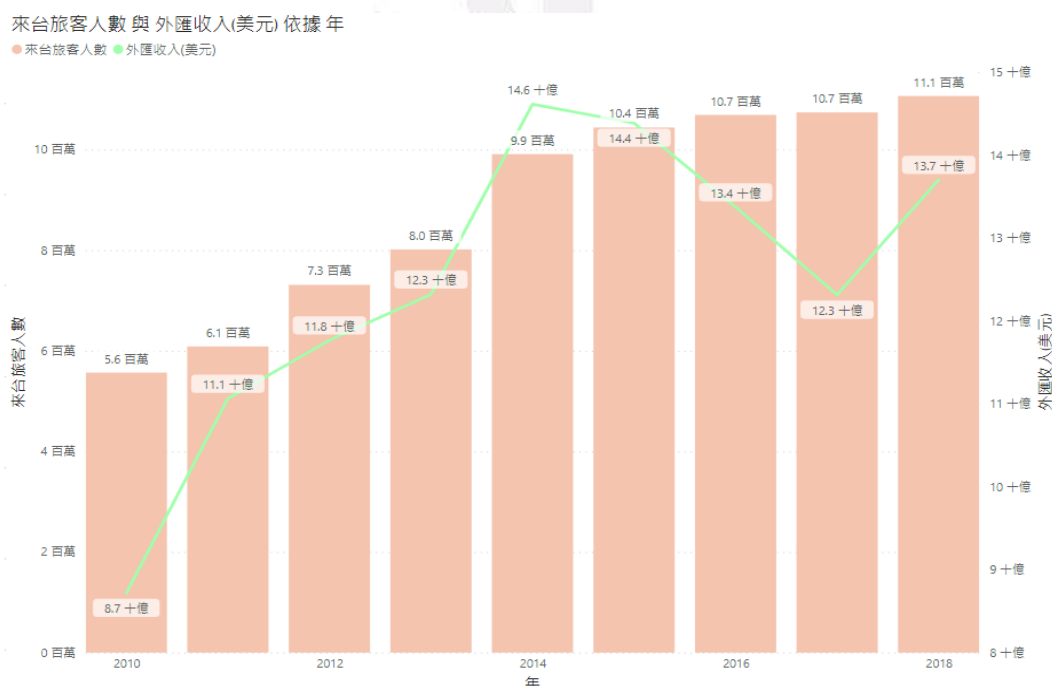
資料來源：交通部觀光局觀光統計資料庫

由表 2 得知，在 2010 年至 2018 年中，8 個遊憩區旅客量最多是在大鵬灣遊憩區，總共有 14125758 人，而旅客量最少的也是位於大鵬灣遊憩區，只有 38526 人。以平均而言，旅客量最多的是在日月潭，平均每年會有 4619378 人去日月潭。比較 8 個遊憩區的標準差，麗寶樂園的標準差是最大，代表它的遊客人數波動程度是最大。

表 2 2010 年至 2018 年主要遊憩區旅遊人數之敘述性統計

	最大值	最小值	平均數	中位數	標準差
九族文化村	2017316	740328	1128138	957891	415753.3
三峽老街	2305940	1849327	2042665	2037600	151871.9
大鵬灣遊憩區	14125758	38526	1763580	461674	207576.1
太魯閣國家公園	2610282	819704	1377966	1022692	681850.6
日月潭	6381636	3203963	4619378	4646006	1055887.0
麗寶樂園	7214438	734073	2394876	997939	2422261.0
駁二藝術特區	4372019	894537	3216770	3632820	1233926.0
鵝鑾鼻公園	1762604	726206	1264689	1364046	337063.5

資料來源：交通部觀光局觀光統計資料庫

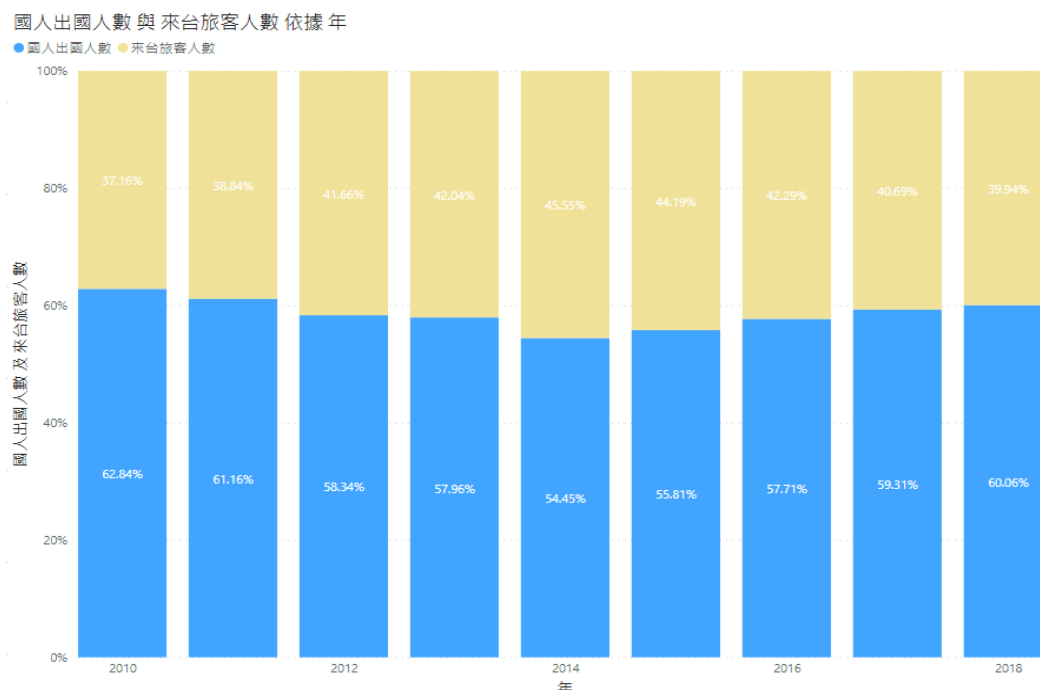


資料來源：交通部觀光局觀光統計資料庫

圖 1 來台旅客人數與外匯收入組合圖

由圖 1 可以觀察出，2010~2018 這個期間來台遊客人數一直只增不減，而外匯收入隨著來台旅客人數的增加而增加，在 2014 年時達到了將近 150 億，是近幾年

來外匯收入最高的時候，而過了 2014 年後，台灣的外匯收入就一直下降，直到 2017 年才又開始慢慢上升。

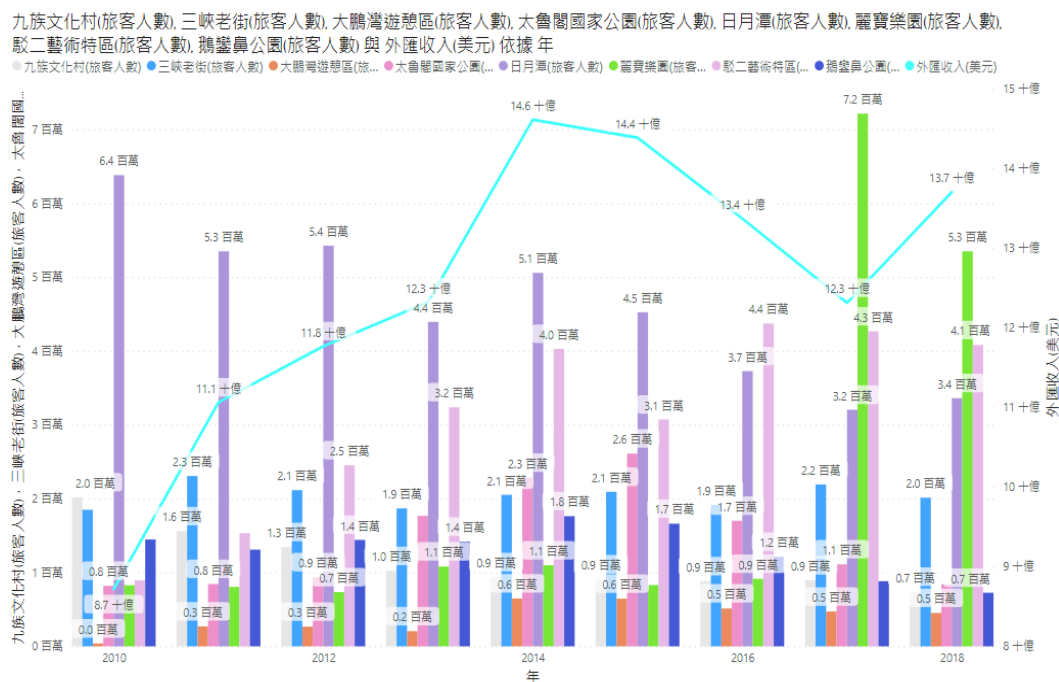


資料來源：交通部觀光局觀光統計資料庫

圖 2 國人出國人數與來台旅客人數

由圖 2 觀察出 2010~2018 這 9 年間，每年的國人出國的人數比例都會高於來台旅客的人數比例，因為隨著國民所得的增加，以及消費水準的提高，國內重視休閒的人口逐漸增加，大家越來越想要往國外去玩，所以每年國人都會安排到國外去旅行，紓解一下這一年來所積累的壓力和疲憊。

## 來台旅遊人數和國人出國人數之時間序列預測分析

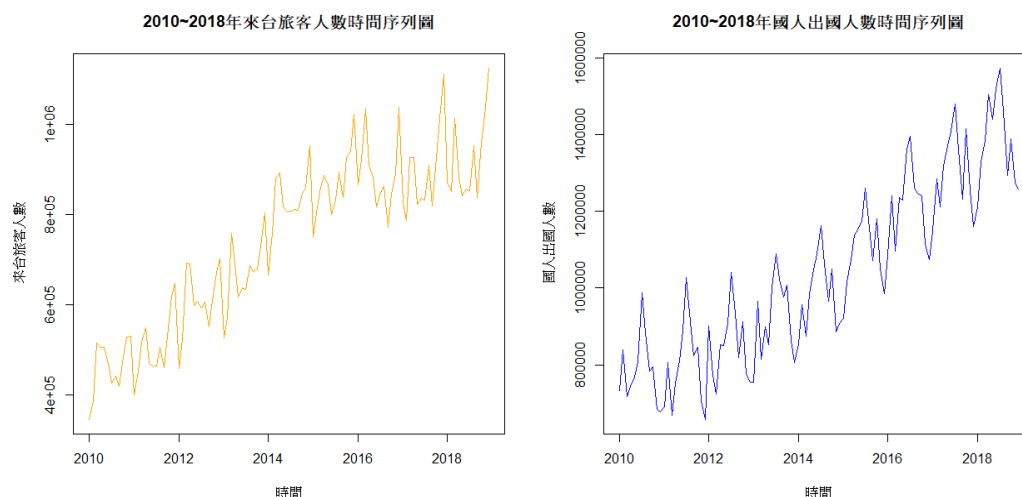


資料來源：交通部觀光局觀光統計資料庫

圖 3 各個遊憩區旅客人數組合圖

圖 3 這些地區是台灣的主要的遊憩區，這些地區每年為台灣帶來了很多外匯，可說是台灣觀光產業的收入來源。由圖 3 可以看出，由圖 3 來台旅客到駁二藝術特區和麗寶樂園在這九年中一直持續增加，直到 2017 年麗寶樂園的人數已經超越其他地區，成為 8 個地區中，觀光人數最多的地方。而日月潭在 2010~2015 年一直是八個地區中觀光人數最多，可是在近幾年來，日月潭的觀光人數一直下降，在 2018 年已經變成 8 個地區的第三名。





資料來源：交通部觀光局觀光統計資料庫

圖 4 來台旅客人數和國人出國人數時間序列圖

圖 4 是 2010~2018 年來台旅客人數時間序列圖和國人出國人數時間序列圖，由來台旅客人數的時間序列圖可以觀察出，旅客人數的資料具有上升的趨勢，但是季節的變動不明顯，無法判別此資料是否具有季節性，所以需要經由在配適模型的過程中，在做進一步的判斷。由國人出國人數的時間序列圖可以觀察出，國人出國人數的資料具有上升的趨勢，而且具有季節性變動，變異數變動程度幾乎一致。

## 第二節 研究動機

過去幾十年來，旅遊不只是一種經濟活動，現今已迅速成為一種大眾的共同需求之一，這種快速發展也導致了國際旅遊的快速崛起，國際旅遊已經成為世界發展最快速的經濟產業之一。隨著各國經濟發展，各國國民對於旅遊的需求增加，觀光產業將會成為各國發展重點，在眾多國家的經濟發展上是不可或缺的角色，因為觀光發展可以增加外匯收入，增加工作機會，以及帶動地方的經濟發展，所以不少國家會致力於推動觀光，希望藉由發展觀光產業來增加就業機會，帶動國家的經濟。

然後隨著時代的變遷，和近年來國民所得的增加，以及生活水準的提高，國內外重視休閒的人口逐漸增加，再加上政府目前實施週休二日，許多家庭人口外出旅遊的比例越來越高，出國觀光旅遊的人數更是直接上升，而來台的旅客人數也是只增不減，所以這種現象造成了台灣觀光業的蓬勃發展，觀光產業將是台灣

未來經濟發展的主力。可是由於各國旅遊業的競爭激烈，除了鄰近台灣的韓國、日本、以及中國大陸之外，東南亞更是一大威脅，因此，如何有效地推動觀光發展提高觀光收入，從而減少失業率，並且讓國家的經濟發展呈現正成長，是台灣目前需要探討的議題。

### 第三節 研究目的

台灣的觀光資源其實非常豐富，包括四通八達的交通運輸便利性、旅遊安全程度高，先天優勢的地理位置加上良好的國際開放度，這些都是有利於我國旅遊市場的發展。再加上政府為了提升我國在國際上的觀光競爭力，政府積極投入並扮演整合中央與地方觀光資源的角色，對內持續優化及精緻化旅遊環境，對外發展自然景觀及人文特色等，透過媒體傳播與海外公關宣傳，傳達臺灣的觀光意向，廣闢來臺客源和市場。所以我想運用在預測分析這堂課所學習到的預測方法來尋找最佳的時間序列模型，並且預測未來一年的台灣的遊客人數，並且找出如何增加台灣遊客人數的對策。

### 第四節 文獻探討及回顧

觀光產業又稱為無煙囪工業，與科技產業被視為是未來兩大明星產業，對就業機會的創造以及外匯的賺取扮演重要的角色。台灣又稱美麗之島，擁有豐富而多樣化的自然資源與人文歷史，極具發展觀光的潛力。隨著國內旅遊人數的增加與國際觀光客來台旅遊人數的逐年成長，我國國際觀光旅館整體住房率也有顯著的提升。所以目前有很多研究是針對台灣的觀光產業為中心，來向外探討有那些因素是會影響台灣的觀光。例如：蔡登茂等人(2008~2009)利用時間序列 ARIMA 預測模式與迴歸分析模式，建構一個國際觀光旅館通用的住房率與整體住房的預測模型架構，同時本研究也以一家國際觀光旅館為研究對象，說明本研究預測模式建構的方法與流程，預測該旅館未來各種客源的住房率與住房總營收，並進行各種客源對該旅館住房總營收的影響分析。其結果發現，國際觀光旅館之營收預測相關問題是國際觀光旅館經營業者所面臨的一個相當重要且複雜的問題，在以往業者大都憑藉著經驗以及對市場的觀察力等定性法進行下一年度的營收預測，不僅過於主觀而且產生的預測誤差都相當大。

近年來政府積極推動國內旅遊市場，由行政院提出的「國內旅遊發展方案」，希望藉由觀光客倍增計畫，整備現有套裝旅遊路線，開發新興觀光景點等策略，發展優質觀光。柳婉郁(2010)針對國內國家級風景特定區之遊客人數建立預測模型，利用 Box-Jenkins 的四個步驟建立 ARMA 預測模型、ARIMA 預測

模型，進一步透過迴歸分析進行實證研究，並以組合預測模式加以比較其預測績效。研究顯示消費者物價指數、國民所得、季節指數及上一期遊客人數、上一年之同期遊客人數對遊客參訪國家級風景特定區有影響。所得增加會增加民眾旅遊之意願，物價指數升高會影響旅遊之消費能力，季節指數顯示第三季高於第一、第二季，顯示第三季為臺灣國家級風景特定區之相對旺季；遊客明顯受到上一期之遊客人數影響，表示資訊傳播對民眾有正面影響。其次，本研究顯示週休二日政策對臺灣遊客人數參訪國家級風景特定區無顯著影響。然後林倉龍(2004)也是利用 Box-Jenkins 的四個步驟建立 ARIMA 預測模型以及利用迴歸分析的最小平方法進行實證研究，針對國內國家風景區區分為北、中、南、東四大區域建立，並以組合預測模式加以比較其預測績效。研究顯示預測組合確實能改善模型的精確度，組合預測融合了兩種預測方法理論的優點，所得到較精確的預測值，本研究顯示北部、中部及南部以 ARIMA 模式求出之預測值有較高的精確度。最後結果也顯示 93 年和 94 年國家風景特定區遊客人次有成長的趨勢，但成長趨勢不大，只有大約 10% 的微幅增長。

回顧歷年來之來台旅客，雖皆以亞洲旅客居多，但有鑒於美、加及歐洲地區亦是來台旅遊之新興市場，同時也為我國在觀光客倍增計畫中努力拓展來台旅遊之目標市場。邱鳳臨、陳奕均(2010)以美國、加拿大和歐洲為研究對象，採用交通部觀光局所公佈的 1971 年 1 月至 2008 年 12 月之旅客入境人數年、季和月資料，分別對此三個地區建立簡單線性迴歸模型(linear trend)、Holt-Winters 趨勢的季節性加法預測模型、自我迴歸自我迴歸移動平均模型 (autoregressive autoregressive moving average)、季節性與非季節性自我迴歸整合移動平均模型 (seasonal-nonsense autoregressive integrated moving average) 和部分整合自我迴歸移動平均模型 (fractionally integrated autoregressive moving average) 等五種模型，進行樣本內的預測，並使用平均絕對百分比誤差和均方根誤差統計量檢驗不同預測模型之預測能力。其結果發現三個來台市場在年資料方面 ARIMA 模型的預測能力最佳；在季和月資料方面，則 ARFIMA 模型的預測能力最好。因此，建議相關單位及旅遊業者從事旅遊預測時可以應用此兩種模式。

交通運輸是旅遊活動形式，也是旅遊的吸引力因素：交通工具的多樣化可豐富活動內容，藉由不同工具進入不同屬性的旅遊地點，增添旅遊活動的色彩與內容，交通工具的體驗有時也是旅遊的目的或吸引因素，交通運輸收入是旅遊業的重要營收之一：除步行與騎自行車外，交通工具的搭乘都必需支付相當的費用，旅客的交通支出通常佔旅遊總預算的 20-40%，旅遊交通運輸不但促進旅

遊業的發展，也是該產業中重要的收益。連聖皓等人(1996)採用時間序列的分析理論，研究臺灣省汽車客運、公民營公車客運、省營汽車客運為研究對象。搜集民國七十年一月至八十二年十二月各月延人公里的月資料為觀察值，作一預測模式然後將預測值與實際值作比較。並進一步預測八十三及八十四年各月數據，藉此瞭解臺灣省汽車客運業的營運狀況及其潛在的問題。分析結果顯示民營汽車客運的延人公里在水平區間發展，但長途客運里程增加、客運人數減少。公民營公車客運的延人公里、行駛里程、客運人數均減少，此乃受自用小客車、機車數量增加所致。而省營汽車客運的延人公里、行駛里程、客運人數也有減少的趨勢，但此只受自用小客車增加的影響並未受機車數量的影響。



## 第二章 研究方法

### 第一節 資料敘述和使用方法說明

基於以上的動機和目的，這次報告我所選取的資料為來台旅客人數、國人出國人數、旅館住用率、以及主要遊憩區的遊客人數按月別統計和按年別統計，分析時間為西元 2010 年 1 月至 2018 年 12 月，此次所選取的資料取自交通部觀光局觀光統計資料庫以及中華民國統計資訊網。此次預計所要分析的方法除了上課所教過的時間序列回歸法(Time series regression)、指數平滑法(Exponential smoothing)、非季節性和季節性的 Box-Jenkins 模型 (Nonseasonal Box-Jenkins models and Seasonal Box-Jenkins models)等方法來進行資料的分析與預測，還會將此資料以數據視覺化的方式，來探討這 9 年來台遊客人數和所選取的遊憩區的旅客人數的變化。



## 第二節 研究流程

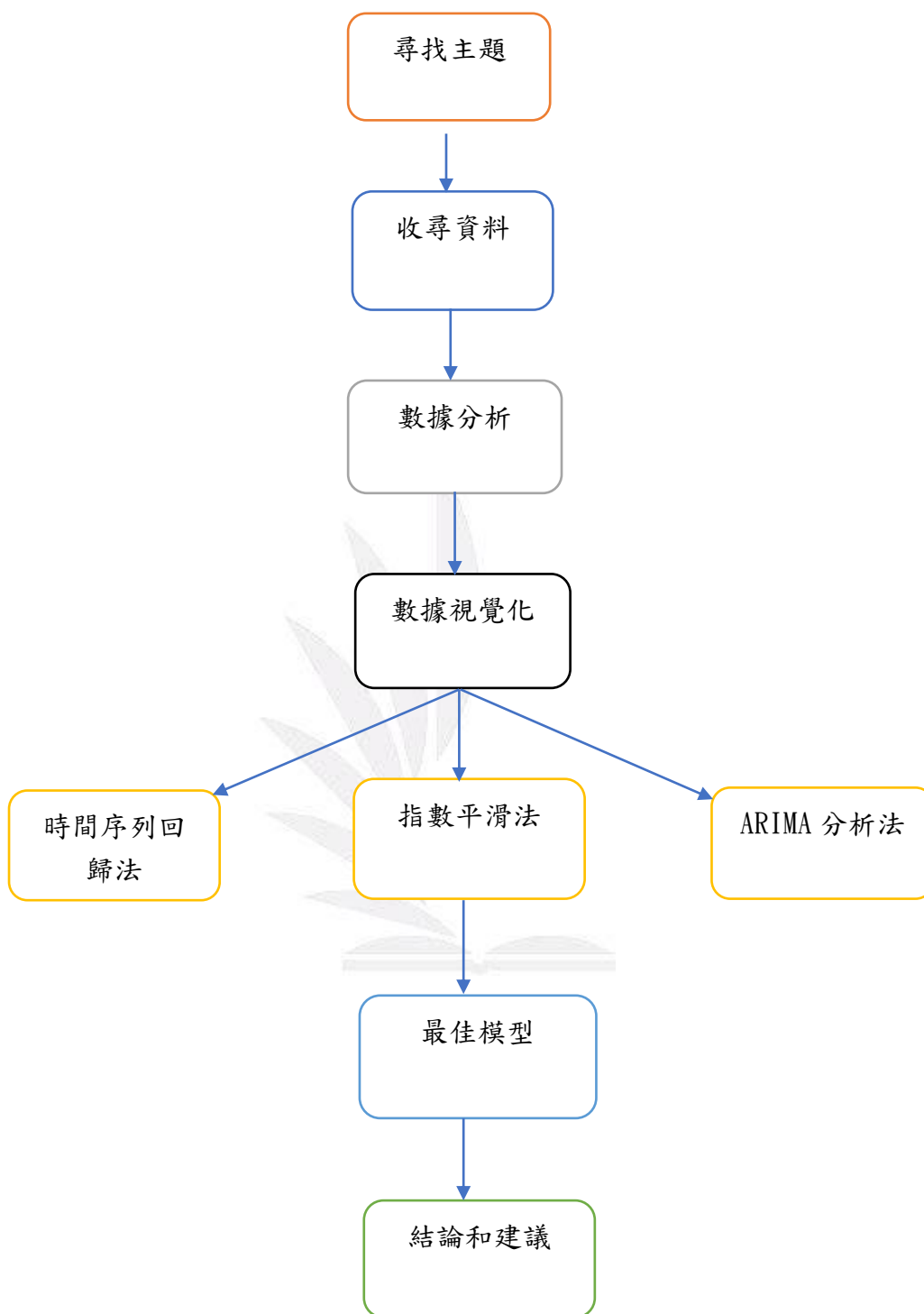


圖 5 分析研究流程圖



### 第三章 研究模型探討

#### 第一節 時間序列回歸法(Time series regression)

時間序列分析是分析時間序列數據的方法，它是使用模型根據先前觀察到的值預測未來值。時間序列分析和迴歸分析的不同在於時間依賴，就是觀察值之間都會存在著相關性，時間相隔越短，相關性越大，如此一來，就不符合迴歸分析當中獨立的基本假設。隨著趨勢的增加或減少，大多數時間序列具有某種形式的季節性趨勢，即特定時間範圍的特定變化。

本次研究的資料是來台的旅客人數和國人出國的人數，由圖 6 可以看出來台旅客的人數是不具有季節變化的，而且隨著時間有長期上升的趨勢，但是上升的趨勢並不是直線上升，所以我將來台的旅客人數配適一個成長曲線模型。

#### 3-1-1 預設模型

由表 3-1 可知， $\beta_1$  為  $0.00776 > 0$ ，所以資料是具有向上升的趨勢，再由圖 6 可以看出它上升的趨勢並不是直線上升，所以我的預設模型為成長曲線模型：

$$y_t = \beta_0 \beta_1^t \varepsilon_t$$

$$\ln(y_t) = \ln(\beta_0) + t \ln(\beta_1) + \ln(\varepsilon_t)$$

t:時間

$y_t$ :在連續時間下，第 t 點的觀察值

表 3-1 成長曲模型的參數估計結果

Parameter Estimates					
Variable	DF	Parameter Estimate	Standard Error	T Value	P-value
Intercept	1	13.05	0.03	538.82	< 0.001
time	1	0.01	0.01	20.12	< 0.001

### 3-1-2 診斷模型

由表 3-2 可以看出，Durbin-Watson 值為 0.872，所以殘差具有一階自相關，Pr < DW 的 p-value 是小於 0.05，代表殘差具有一階正自我相關，Pr > DW 的 p-value 是大於 0.05，代表殘差不具有一階負自我相關，所以這個成長曲線模型是具有一階正自我相關，所以必須修正這個模型。

表 3-2 成長曲線模型自相關檢定表

Durbin-Waston D	0.872
Pr < DW	< 0.001
Pr > DW	1
Number of Observations	108
1st Order Autocorrelation	0.535

### 3-1-3 修正模型

因為預設模型的殘差具有一階正自我相關，所以我把殘差去配 AR(1)模型，修正後的模型：

$$y_t = \beta_0 \beta_1^t \varepsilon_t$$

$$\ln(y_t) = \ln(\beta_0) + t \ln(\beta_1) + \ln(\varepsilon_t)$$

$$\text{Where } \varepsilon_t = \varphi_1 \varepsilon_{t-1} + a_t$$

t:時間

$y_t$ :在連續時間下，第 t 點的觀察值

### 3-1-4 判斷修正的模型是否合適

由表 3-3 可以看出，殘差自相關檢定的 p-value 都是小於 0.05，代表修正過後的模型還是具有一階自相關，代表來台旅客人數的資料是不適合用成長曲線模型台建模。

由圖 6 可以看出，國人出國的人數是具有季節性，而且隨著時間有長期上升的趨勢，所以我將國人的出國人數配是一個具有趨勢項和季節項的時間序列迴歸模型。



表 3-3 修正後成長曲線模型自相關檢定表

殘差的自相關檢查									
至滯後	卡方	DF	p-value	自相關					
6	24.85	5	0.0001	0.019	-0.248	0.023	0.364	0.111	-0.105
12	112.86	11	< 0.001	0.077	0.37	-0.039	-0.261	-0.035	0.711
18	135.64	17	< 0.001	-0.013	-0.29	-0.025	0.271	0.048	-0.13
24	195.34	23	< 0.001	0.007	0.281	-0.06	-0.233	-0.093	0.531

### 3-2-1 預設模型

由表 3-4 可知， $\beta_1$  是 6708.21 > 0，所以資料是具有上升的直線趨勢，再由圖 6 可以看出它上升的趨勢是直線上升，所以我的預設模型為具有趨勢項和季節項的時間序列迴歸模型：

$$y_t = TR_t + SN_t + \varepsilon_t = \beta_0 + \beta_1 t + \beta_2 M_1 + \beta_3 M_2 + \dots + \beta_{12} M_{11} + \varepsilon_t$$

t:時間

$y_t$ :在連續時間下，第 t 點的觀察值

$M_1, M_2, \dots, M_{11}$  為季節虛擬變數

For example  $M_{11} = \begin{cases} 1, & \text{if period is November} \\ 0, & \text{otherwise} \end{cases}$

表 3-4 時間序列迴歸模型參數估計表

Parameter Estimates					
Variable	DF	Parameter Estimate	Standard Error	T Value	P-value
Intercept	1	517363	21808	23.72	< 0.001
time	1	6708.21	177.35	37.82	< 0.001
M1	1	76266	26991	2.83	0.005
M2	1	170489	26979	6.32	0.0009
M3	1	92315	26968	3.42	< 0.001
M4	1	181760	26958	6.74	< 0.001
M5	1	185130	26949	6.87	< 0.001
M6	1	250226	26942	9.29	< 0.001
M7	1	338358	26935	12.56	< 0.001
M8	1	224208	26930	8.33	< 0.001
M9	1	123966	26926	4.6	< 0.001
M10	1	186590	26923	6.93	< 0.001
M11	1	42648	26921	1.58	< 0.001

### 3-2-2 診斷模型

由表 3-5 可以看出，Durbin-Watson 值為 0.927，所以殘差具有一階自相關，Pr < DW 的 p-value 是小於 0.05，代表殘差具有一階正自我相關，Pr > DW 的 p-value 是大於 0.05，代表殘差不具有一階負自我相關，所以這個時間序列迴歸模型是具有一階正自我相關，所以必須修正這個模型。

表 3-5 時間序列迴歸模型自相關檢定表

Durbin-Waston D	0.927
Pr < DW	< 0.001
Pr > DW	1
Number of Observations	108
1st Order Autocorrelation	0.508

### 3-2-3 修正模型

因為預設模型的殘差具有一階正自我相關，所以我把殘差去配 AR(1)模型，修正後的模型：

$$y_t = TR_t + SN_t + \varepsilon_t = \beta_0 + \beta_1 t + \beta_2 M_1 + \beta_3 M_2 + \dots + \beta_{12} M_{11} + \varepsilon_t$$

Where  $\varepsilon_t = \varphi_1 \varepsilon_{t-1} + a_t$

t:時間

$y_t$ :在連續時間下，第 t 點的觀察值

$M_1, M_2, \dots, M_{11}$  為季節虛擬變數

For example  $M_{11} = \begin{cases} 1, & \text{if period is November} \\ 0, & \text{otherwise} \end{cases}$

#### 3-1-4 判斷修正的模型是否合適

雖然在滯後 18 的 P-value 是小於 0.05，但是整體上來說，整個時間序列迴歸模型自相關檢定的 p-value 是大於 0.05，此時間序列迴歸模型是不具有一階自相關，表示國人出國人數用此時間序列迴歸模型來建模是合適。

表 3-6 修正後時間序列迴歸模型自相關檢定表

殘差的自相關檢查									
至滯後	卡方	DF	p-value	自相關					
6	3.51	5	0.6217	-0.098	0.07	-0.012	0.069	0.082	0.069
12	17.66	11	0.0897	0.125	0.058	0.02	0.114	0.246	-0.154
18	31.41	17	0.0178	0.271	-0.022	-0.006	-0.008	0.012	0.184
24	34.01	23	0.065	-0.123	0.052	0.031	-0.01	0.01	0.022

表 3-7 時間序列回歸法真實值與預測值

時間	真實值	預測值
2019/01	1412220	600337.9
2019/02	603211	701269.3
2019/03	101217	629802.9
2019/04	11168	725956.6
2019/05	17611	736034.3
2019/06	18249	807839.3
2019/07	28106	902679.4
2019/08	37786	795237.4
2019/09	31622	701703.5
2019/10	30390	771035.8
2019/11	23666	633801.7
2019/12	20318	597862

表 3-8 時間序列回歸評估指標表

Model	MAE	MAPE	MSE
AR(1)模型	657646.7	25.41296	470056497342

## 第二節 指數平滑法(Exponential smoothing)

指數平滑法它是利用過去時間的加權平均值來平滑資料的方法，它是在移動平均法的基礎上發展出來的一種時間預測方法，它是透過計算指數平滑值，配合一定的時間序列模型來預測。時間序列是沒有趨勢或是季節性，就可以利用 Simple Exponential Smoothing 來配適，時間序列是具有趨勢但無季節性的資料，就用 Holt's Trend Corrected Exponential Smoothing 來配適，時間序列是具有趨勢及固定的季節變異，就用 Additive Holt-Winters Method 來配適，時間序列是具有趨勢及非固定的季節變異，就用 Multiplicative Holt-Winters Method 來配適。

由圖 6 的旅客人數時間序列圖可以看出，旅客人數的資料具有上升的趨勢，但是季節的變動不明顯，所以我們先決定用 Holt's Trend Corrected Exponential Smoothing 來配適。

**預設模型:**

$$l_T = \alpha y_t + (1 - \alpha)[l_{T-1} + b_{T-1}]$$

$$b_T = \gamma[l_T - l_{T-1}] + (1 - \gamma)b_{T-1}$$

$$\hat{y}_{t+\tau}(T) = l_T + \tau b_T$$

Smoothing parameters:

$$\alpha = 0.1978$$

$$\beta = 0.0183$$

表 3-9 來台旅客人數模型比較表

分析方法	MAE	MAPE	MSE
Holt's Trend Corrected Exponential Smoothing	60223.49	8.398772	5435390370
Additive Holt-Winters Method	27259.5	3.773783	1263706517
Multiplicative Holt-Winters Method	29587.17	4.108705	1532250387
各準則下的最佳模型	Additive Holt-Winters Method	Additive Holt-Winters Method	Additive Holt-Winters Method

經由 MAE、MAPE、MSE 三個評估準則可以判斷出，其實來台旅客人數最佳的指數平滑模型為 Additive Holt-Winters Method。

**最佳指數平滑模型:**

$$l_T = \alpha y_t + (1 - \alpha)[l_{T-1} + b_{T-1}]$$

$$b_T = \gamma[l_T - l_{T-1}] + (1 - \gamma)b_{T-1}$$

$$sn_T = \delta(y_T - l_T) + (1 - \delta)sn_{T-L}$$

$$\hat{y}_{t+\tau}(T) = l_T + \tau b_T + sn_{T+\tau-L}$$

Smoothing parameters:

$$\alpha = 0.6197$$

來台旅遊人數和國人出國人數之時間序列預測分析

$\beta=0.004$

$\gamma=0.0183$

表 3-10 來台旅客人數 Additive Holt-Winters Method 真實值與預測值

時間	真實值	預測值
2019/01	1412220	921347.2
2019/02	603211	969045.5
2019/03	101217	1096183.8
2019/04	11168	1066289
2019/05	17611	1004838.5
2019/06	18249	992883.1
2019/07	28106	1003367.3
2019/08	37786	1039520.5
2019/09	31622	993852.6
2019/10	30390	1059389.1
2019/11	23666	1112539.2
2019/12	20318	1190315

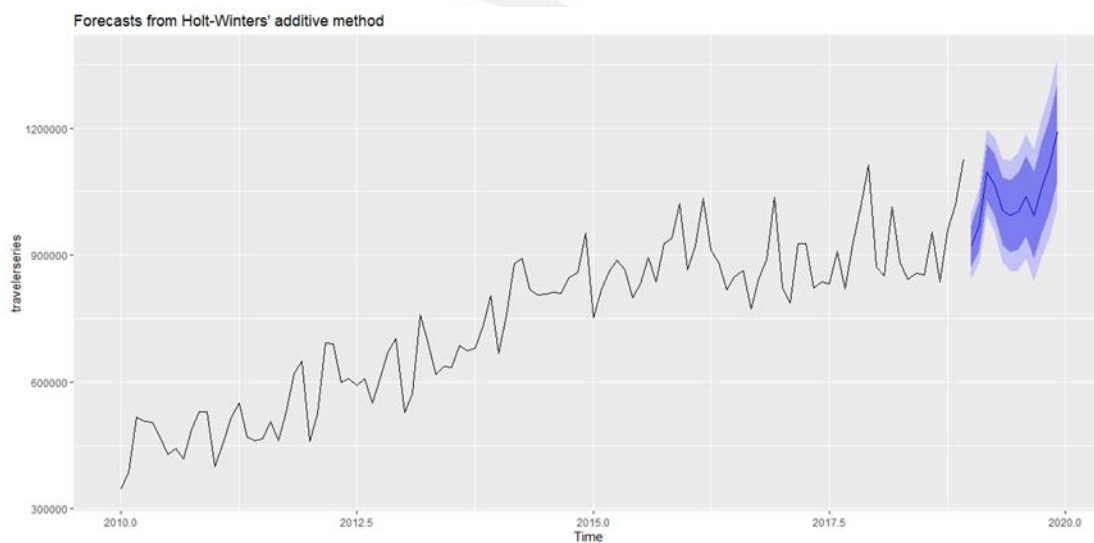


圖 3-1 來台旅客人數 Additive Holt-Winters Method 預測圖

由圖 6 的國人的出國人數時間序列圖可以看出，國人出國人數具有上升的趨勢，而且具有季節變動，變異數變動的程趨於一致，所以我們決定用 Additive Holt-Winters Method 來配適。

**預設模型：**

$$l_T = \alpha y_t + (1 - \alpha)[l_{T-1} + b_{T-1}]$$

$$b_T = \gamma[l_T - l_{T-1}] + (1 - \gamma)b_{T-1}$$

$$sn_T = \delta(y_T - l_T) + (1 - \delta)sn_{T-L}$$

$$\hat{y}_{t+\tau}(T) = l_T + \tau b_T + sn_{T+\tau-L}$$

Smoothing parameters:

alpha=0.0886

beta=0.015

gamma=0.01831

表 3-11 國人出國人數模型比較表

分析方法	MAE	MAPE	MSE
Holt's Trend Corrected Exponential Smoothing	84820.09	8.520319	9482537206
Additive Holt-Winters Method	34192.13	3.431912	2051258376
Multiplicative Holt-Winters Method	39008.59	3.839073	2403659476
各準則下的最佳模型	Additive Holt-Winters Method	Additive Holt-Winters Method	Additive Holt-Winters Method

經由 MAE、MAPE、MSE 三個評估準則可以判斷出，國人出國人數最佳的指數平滑模型為 Additive Holt-Winters Method。

表 3-12 國人出國人數 Additive Holt-Winters Method 真實值與預測值

時間	真實值	預測值
2019/01	1412220	1334460
2019/02	603211	1433239
2019/03	101217	1366767
2019/04	11168	1470502
2019/05	17611	1477327
2019/06	18249	1551619
2019/07	28106	1636735
2019/08	37786	1529649
2019/09	31622	1443883
2019/10	30390	1506973
2019/11	23666	1364062
2019/12	20318	1322984

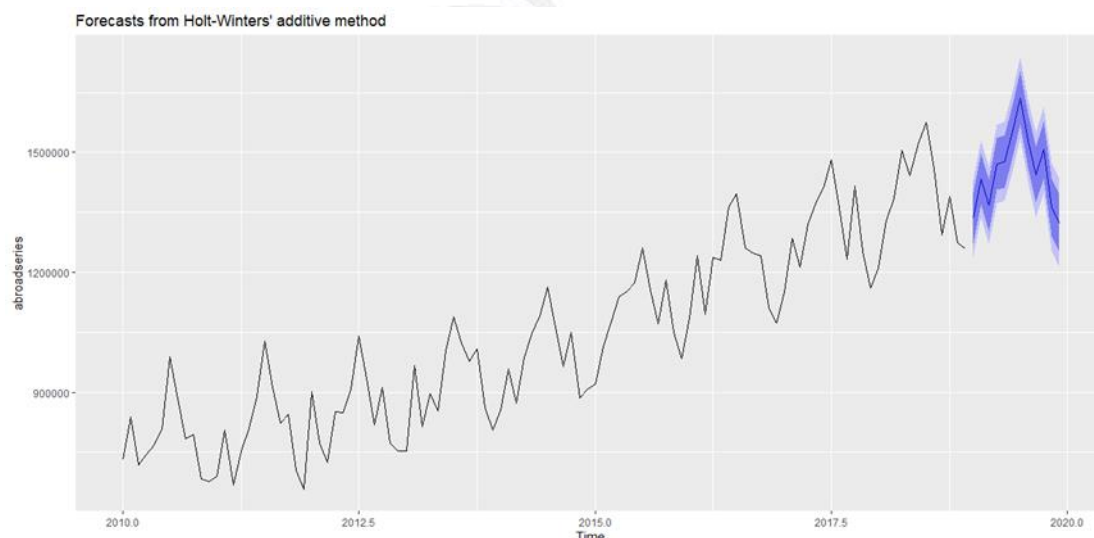


圖 3-2 國人出國人數 Additive Holt-Winters Method 預測圖



### 第三節 ARIMA 分析法

ARIMA 模型(Autoregressive Integrated Moving Average model)是由 Box and Jenkin 所提出的差分整合移動平均自迴歸模型，又稱為整合移動平均自回歸模型，ARIMA 的基本構想為將預測的時間序列資料視為一個隨機序列，用一定的模型來近似描述這個序列，模型一旦確認後就可以從時間序列的過去值和現在值來預測來預測未來值。

由圖 6 的旅客人數時間序列圖可以看出，其實季節的變動不明顯，但由圖 3-3 可以得知其實旅客人數資料是具有固定的季節變動。旅客人數的資料具有上升的趨勢，代表資料還未達到平穩，由圖 3-3 的 ACF 圖可以看出，資料的 ACF 為 Dies down slowly，所以確定資料是不平穩，再發現 lag12 和 lag24 有特別突出的現象，所以接下來的步驟就是對資料做一次差分或是季節差分。

將資料做了一次差分、季節差分、和同時做季節差分和一次差分後，互相比較過後，發現同時做季節差分和一次差分後的 ACF(圖 3-4)收斂得比較快，所以最後認為同時對資料做了季節差分和一次差分最適當。做了一次差分和季節差分過後，此時間序列資料已經達到平穩，接下來就是進一步配適模型。

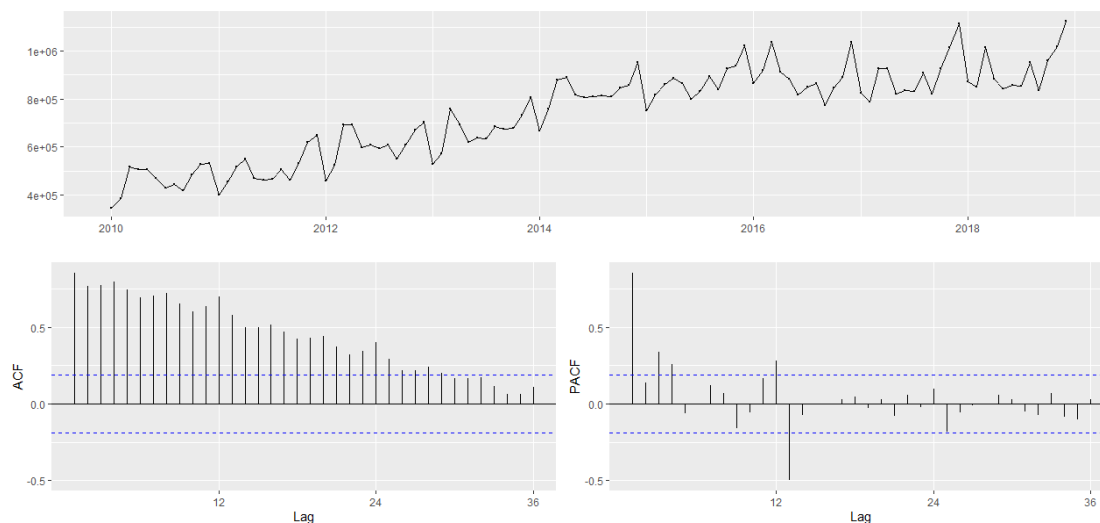


圖 3-3 來台旅客人數原始資料的時間序列圖、ACF 圖、PACF 圖

來台旅遊人數和國人出國人數之時間序列預測分析

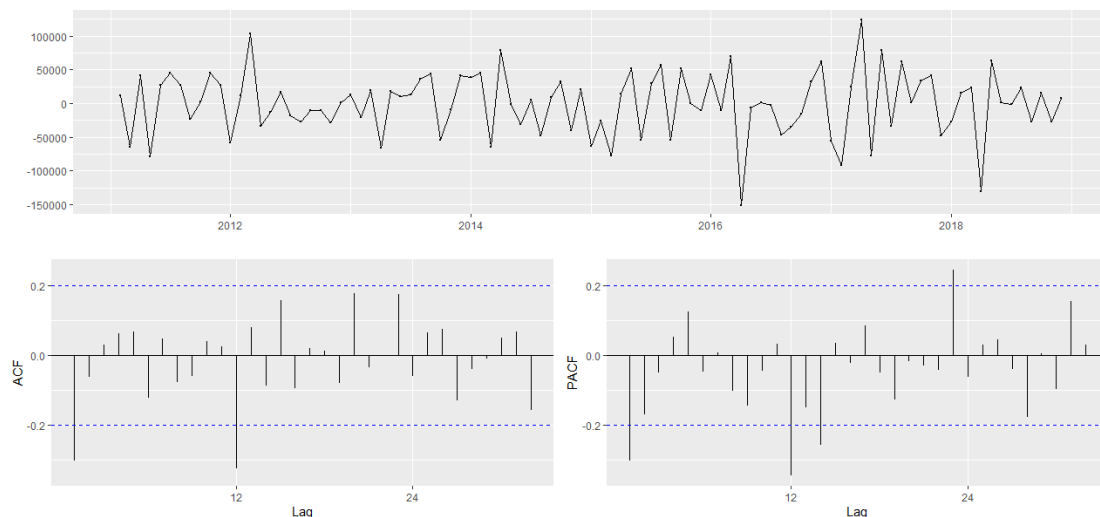


圖 3-4 來台旅客人數資料經過一次差分及季節差分後之時間序列圖、ACF 圖、PACF 圖

由經過一次差分和季節差分後的 ACF 和 PACF 圖(圖 3-4)得知，ACF 呈現出較好的 cut off，PACF 呈現的是 dies down，所以要配適 MA 模式，最終我們決定對資料配適 ARIMA(0,1,1)(0,1,1)。

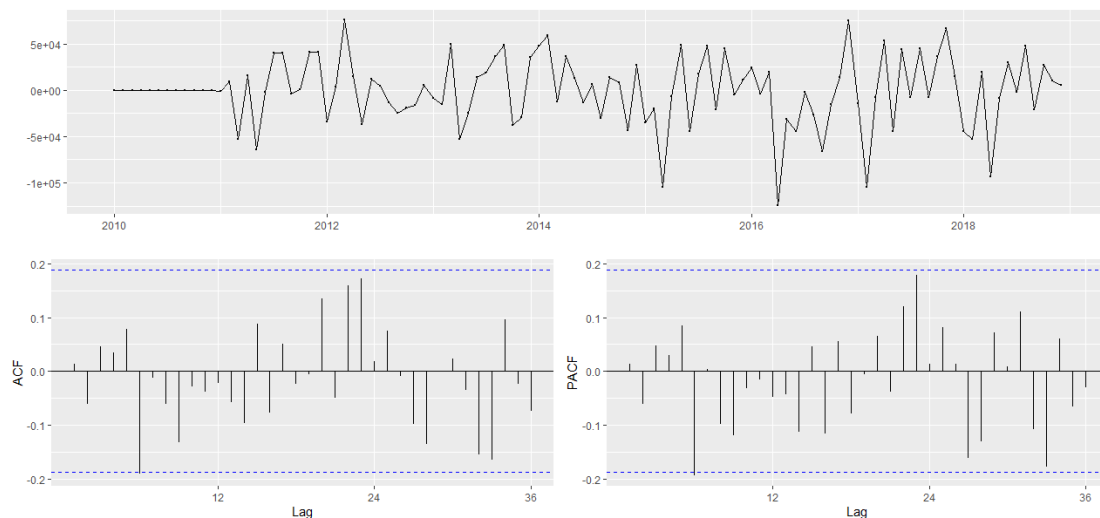


圖 3-5 ARIMA(0,1,1)(0,1,1) 的時間序列圖、ACF 圖、PACF 圖

表 3-13 ARIMA(0,1,1)(0,1,1)評估指標表

Model	MAE	MAPE	MSE
ARIMA(0,1,1)(0,1,1)	27080.45	3.574827	1360591013

配完模型後，我們將對我們的模型進行自我相關檢定，我們將用 White noise、單根檢定、Ljung-Box 來對殘差進行檢驗，檢驗我們模型是否具有自我相關。

我們先建立各個檢定的虛無假設和對立假設

1. White noise 檢定

$H_0$ : 具有 White noise

$H_1$ : 不具有 White noise

檢定規則: 若 p-value 小於 0.05，則拒絕  $H_0$ ，表示模型的殘差不具有 White noise。

2. 單根檢定

$H_0$ : 時間序列不平穩

$H_1$ : 時間序列平穩

檢定規則: 若 p-value 小於 0.05，則拒絕  $H_0$ ，表示時間序列達到了平穩。

3. Ljung-Box 檢定

$H_0: \rho_1 = \rho_2 = \dots = \rho_n = 0$

$H_1$ : 至少一個不為 0

檢定規則: 若 p-value 小於 0.05，則拒絕  $H_0$ ，表示時間序列達到了平穩。

表 3-14 ARIMA(0,1,1)(0,1,1)單根檢定表

R 程式碼方法	p-value	是否拒絕 $H_0$	是否平穩
ur.kpss	大於 0.05	不拒絕 $H_0$	平穩
ur.df	小於 0.05	拒絕 $H_0$	平穩
ur.pp	小於 0.05	拒絕 $H_0$	平穩

表 3-15 ARIMA(0,1,1)(0,1,1)Ljung-Box 檢定表

檢定方法	p-value	是否拒絕	是否具有白噪音
Ljung-Box	0.5345>0.05	不拒絕 $H_0$	有白噪音

表 3-16 ARIMA(0,1,1)(0,1,1)真實值和預測值

時間	真實值	預測值
2019/01	1412220	908790.2
2019/02	603211	904878.6
2019/03	101217	104442.8
2019/04	11168	967748.3
2019/05	17611	911702
2019/06	18249	909928
2019/07	28106	913057.8
2019/08	37786	988260.1
2019/09	31622	892367.7
2019/10	30390	997107.9
2019/11	23666	1056730.9
2019/12	20318	1162142.4

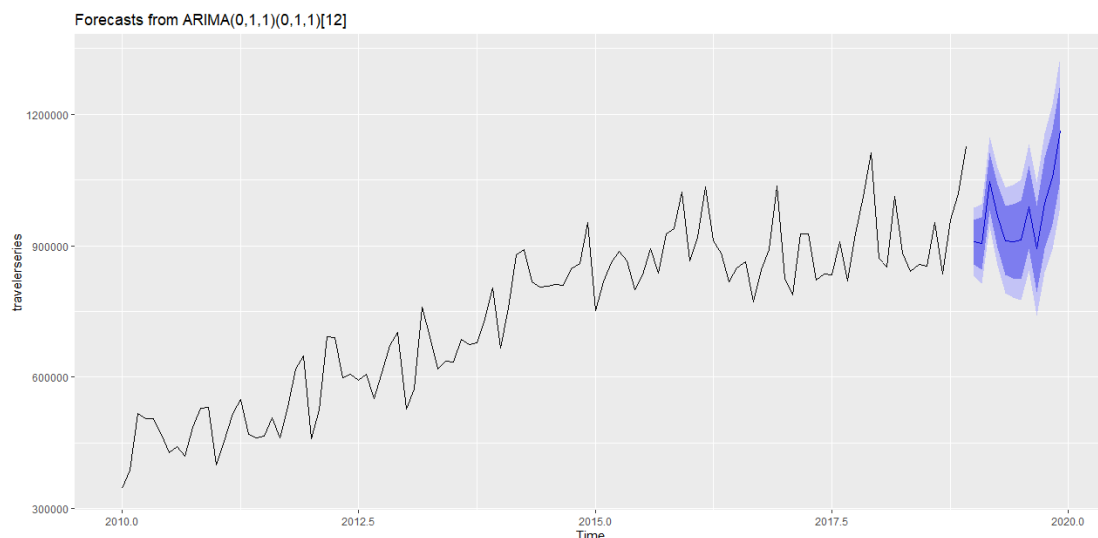


圖 3-6 ARIMA(0,1,1)(0,1,1)預測圖

由圖 6 的國人出國人數時間序列圖可以看出，其實季節的變動是固定的，但由圖 3-7 可以再次確認國人出國人數資料是具有固定的季節變動。旅客人數的資料具有上升的趨勢，代表資料還未達到平穩，由圖 3-7 的 ACF 圖可以看出，資料的 ACF 為 Dies down slowly，所以確定資料是不平穩，再發現 lag12 和 lag24 有特別突出的現象，所以接下來的步驟就是對資料做一次差分或是季節差分。

將資料做了一次差分、季節差分、和同時做季節差分和一次差分後，互相比較過後，發現同時做季節差分和一次差分後的 ACF(圖 3-8)收斂得比較快，所以最後認為同時對資料做了季節差分和一次差分最適當。做了一次差分和季節差分過後，此時間序列資料已經達到平穩，接下來就是進一步配適模型。

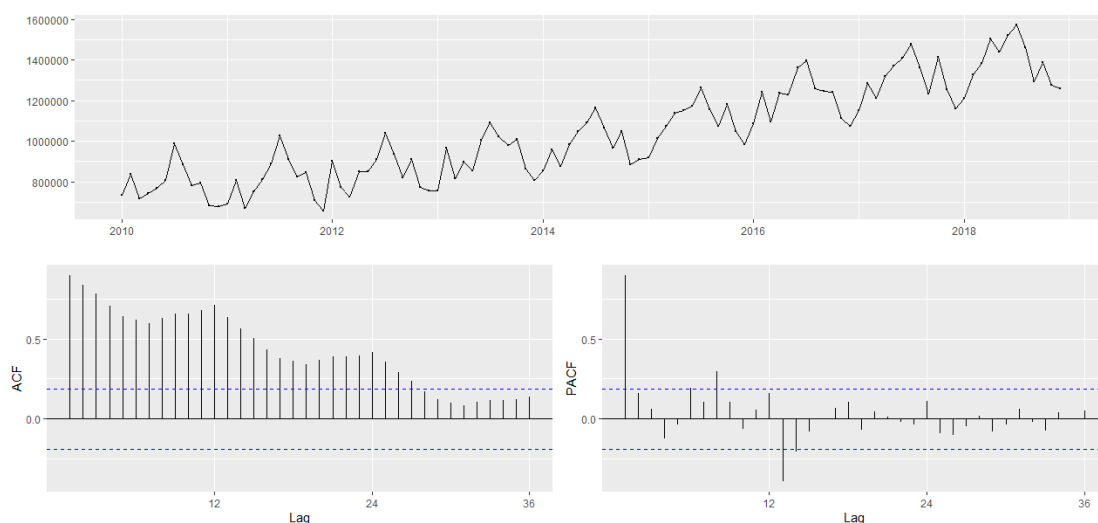


圖 3-7 國人出國人數原始資料的時間序列圖、ACF 圖、PACF 圖

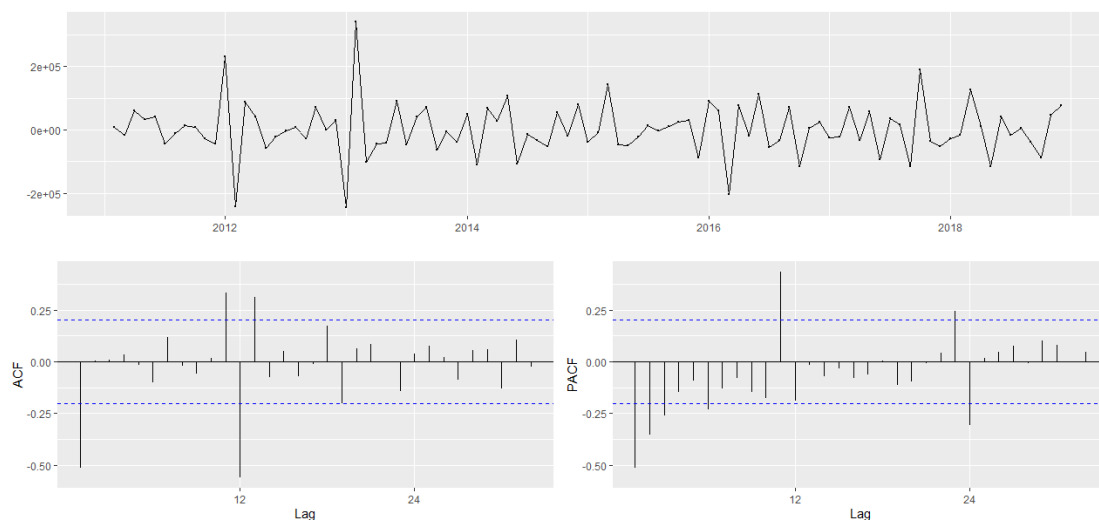


圖 3-8 國人出國人數資料經過一次差分及季節差分後之時間序列圖、ACF 圖、PACF 圖

由經過一次差分和季節差分後的 ACF 和 PACF 圖(圖 3-9)得知，ACF 呈現出較好的 cut off，PACF 呈現的是 dies down，所以要配適 MA 模式，由 PACF 圖可以得知，在 lag12 和 lag24 有較明顯的不顯著，所以最終我們決定對資料配適 ARIMA(0,1,1)(0,1,1)，然後透過 R 程式碼找到的最佳模型為 ARIMA(2,1,2)(1,1,1)，所以以下我們將比較兩個模型：

表 3-17 ARIMA(0,1,1)(0,1,1)和 ARIMA(2,1,2)(1,1,1)的比較表

分析方法	MAE	MAPE	MSE
ARIMA(0,1,1)(0,1,1)	33871.76	3.28957	2257919459
ARIMA(2,1,2)(1,1,1)	31873.89	3.097566	1925763328
各準則下的最佳模型	ARIMA(2,1,2)(1,1,1)	ARIMA(2,1,2)(1,1,1)	ARIMA(2,1,2)(1,1,1)

由 MAE、MAPE、MSE 三個評估準則判斷，最佳的 ARIMA 模型為 ARIMA(2,1,2)(1,1,1)。

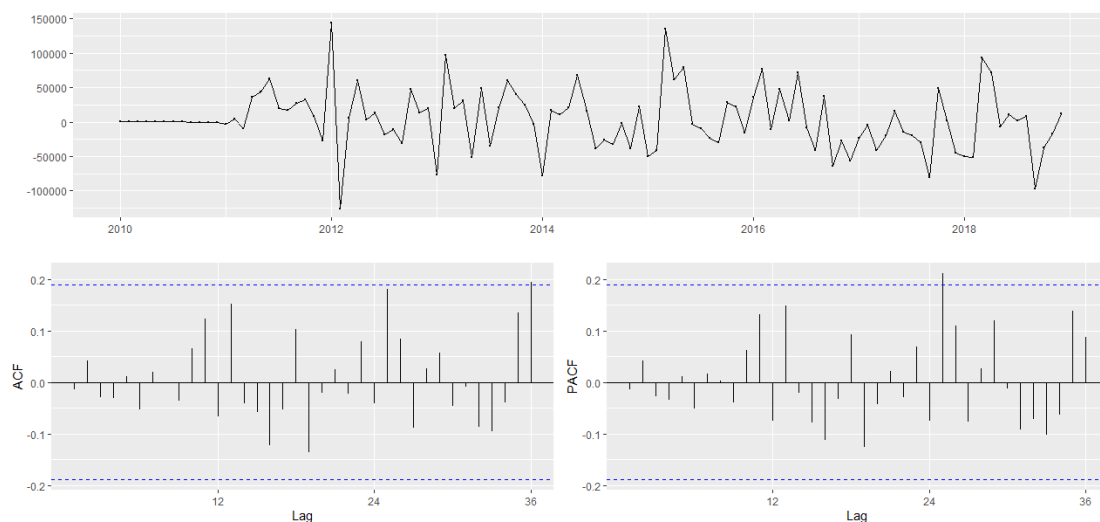


圖 3-9 ARIMA(2,1,2)(1,1,1) 的時間序列圖、ACF 圖、PACF 圖

配完模型後，我們將對我們的模型進行自我相關檢定，我們將用 White noise、單根檢定、Ljung-Box 來對殘差進行檢驗，檢驗我們模型是否具有自我相關。

表 3-18 ARIMA(2,1,2)(1,1,1) 單根檢定表

R 程式碼方法	p-value	是否拒絕 $H_0$	是否平穩
ur.kpss	大於 0.05	不拒絕 $H_0$	平穩
ur.df	小於 0.05	拒絕 $H_0$	平穩
ur.pp	小於 0.05	拒絕 $H_0$	平穩

表 3-19 ARIMA(2,1,2)(1,1,1) Ljung-Box 檢定表

分析方法	p-value	是否拒絕 $H_0$	是否具有白噪音
Ljung-Box	0.6187	不拒絕 $H_0$	具有白噪音

表 3-20 ARIMA(2,1,2)(1,1,1)真實值和預測值

時間	真實值	預測值
2019/01	1412220	1321449
2019/02	603211	1445021
2019/03	101217	1386183
2019/04	11168	1493384
2019/05	17611	1504484
2019/06	18249	1569128
2019/07	28106	1630187
2019/08	37786	1510151
2019/09	31622	1405451
2019/10	30390	1512511
2019/11	23666	1367210
2019/12	20318	1308018

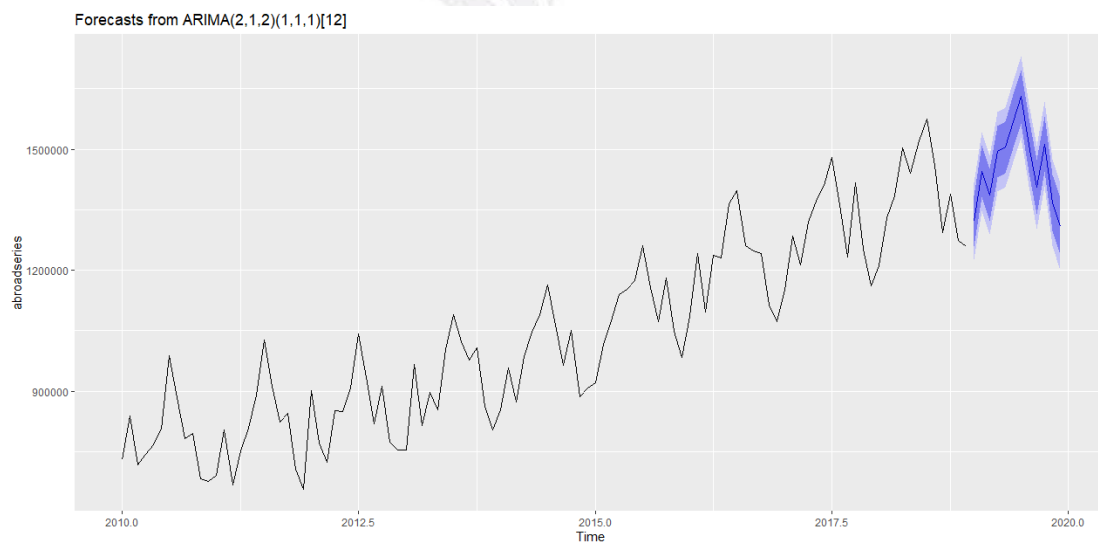


圖 3-10 ARIMA(2,1,2)(1,1,1)的預測圖



#### 第四節 最佳模型

我們利用時間序列迴歸法、指數平滑法、以及 ARIMA 之分析配適預測模式，再利用 MAE、MSE 和 MAPE 三個準則來判斷哪一個模型為最佳預測模型，MAE、MSE 和 MAPE 三個準則為越小，代表該預測模型越佳。我將來台旅客人數和國人出國人數的模型評估呈現於表 3-21 和 3-22。由表 3-21 可以觀察出來，在三個準則下，來台旅客人數用指數平滑法配適模型，要比 ARIMA 分析法來的好。由表 3-22 可以觀察出，ARIMA 分析法在所有準則下的表現都是最佳的，代表國人出國人數用 ARIMA 分析法來配適預測模型，預測出來的結果，比時間序列迴歸法和指數平滑法還要好。綜合以上敘述，來台旅客人數我會選擇用指數平滑法來配適預測模型，而國人出國人數我會選擇用 ARIMA 分析法來配適預測模型。

表 3-21 來台旅客人數模型評估表

分析方法	MAE	MAPE	MSE
指數平滑法	27259.5	3.773783	1263706517
ARIMA 分析法	31873.89	3.097566	1925763328
各準則下的最佳模型	指數平滑法	ARIMA 分析法	指數平滑法

表 3-22 國人出國人數模型評估表

分析方法	MAE	MAPE	MSE
時間序列迴歸法	657646.7	25.41296	470056497342
指數平滑法	34192.13	3.431912	2051258376
ARIMA 分析法	31873.89	3.097566	1925763328
各準則下的最佳模型	ARIMA 分析法	ARIMA 分析法	ARIMA 分析法

## 第四章 結論與建議

在此報告中，我們分別使用時間序列回歸法、指數平滑法、以及 ARIMA 分析法來分析來台旅客人數和國人出國人數的時間序列資料，然後配適預測模型，並且藉由 MAE(平均絕對誤差)、MAPE(平均絕對百分誤差)、MSE(均方誤差)這三個評估準則來挑選出最佳預測模型，以來台旅客人數而言，指數平滑法在三個準則下是表現最良好的，然後將圖 3-1(指數平滑法的來台旅客人數預測圖)和圖 3-6 (ARIMA 分析法的來台旅客人數的預測圖)做比較，可以發現指數平滑法的來台旅客人數預測圖所呈現的預測趨勢跟真實的趨勢比較相符合，所以我們才將指數平滑法所配適的預測模型當做來台旅客人數的最佳模型。而以國人出國人數而言，ARIMA 分析法在三個準則下所呈現的效果是最佳的，然後由圖 3-10 (ARIMA 分析法的來台旅客人數的預測圖)可以觀察的出，預測出來的趨勢跟真實的情況的趨勢大約一致，所以我才將 ARIMA 分析法所配適的預測模型當做國人出國人數的最佳模型。

我們可以從各個預測模型和預測模型的預測圖可以得知，不管是來台旅客人數還國人出國人數依舊還是會持續上升，從來台旅客人數的最佳模型的預測情況可以看出除了在 4 月、5 月、6 月、9 月會有小幅度的下降之外，其餘月份都是穩定上升的。然後從國人出國人數的最佳模型的預測情況可以觀察的出，也是有少許幾個月又微幅下將之外，其餘月份也是有上升的趨勢。

從來台旅客人數的預測情況來看，代表會有越來越多人來台灣遊玩，而為了讓台灣的觀光產業能夠永續發展，除了觀光客量的增加外，更要兼顧觀光品質的提升，近年來隨著觀光休閒產業的蓬勃發展，使得觀光遊憩活動成為國際旅客來台遊玩及國人生活中的一部分，觀光產業需要有良好的規劃，讓來臺旅客可以輕易的親近旅遊目的地。國人應加強規劃與設置觀光產業的軟、硬設施，為迎接更多的國外旅客做準備。

從國人出國人數的預測情況來看，也代表會有越來越多台灣國人想到國外去遊玩，除了跟近年來國民所得的增加，以及生活水準的提高有關之外，或許也跟台灣人想要切換一個旅遊環境有關，台灣人從小到大就時常在台灣這片土地上旅行遊玩，所以已經對台灣的旅遊環境已經非常熟悉，所以這種情況導致越來越多台灣人想要到國外去玩，去一個跟自己生活完全不一樣的陌生環境，感受一個與平時生活環境截然不同的風俗和環境。

## 參考文獻

1. 連聖皓、朱宜寧、謝邦昌。「臺灣省汽車客運量之時間序列預測模式」。輔仁管理評論，台北市：輔仁大學統計學系、台北市政府主計處。
2. 蔡登茂、陳建助、鄭如嵐(2008~2009)「國際觀光旅館整體住房營收預測模式之研究-以某國際觀光旅館為例」，行政院國家科學委員會專題研究計畫成果報告
3. 柳婉郁(2010)「我國國家級風景特定區觀光遊憩人次預測模式之研究」，真理大學自然資源應用學系
4. 林倉龍(2004)「國家風景特定區遊客人次預測之研究」朝陽科技大學休閒事業管理系碩士論文
5. 邱鳳臨、陳奕均(2010)「歐美地區對台灣之旅遊需求預測」。國家發展研究所
6. 中華民國統計資訊網：  
<https://www1.stat.gov.tw/ct.asp?xItem=15466&CtNode=4763&mp=3>
7. 交通部觀光局觀光統計資料庫：<https://stat.taiwan.net.tw/introduction>