

基于Auto-Regressive的河北省旅游接待人数预测研究

聂冉冉¹, 李志新^{1*}, 李志国²

¹河北工程大学数理科学与工程学院, 河北 邯郸

²河北工业大学理学院, 天津

Email: niezairan@163.com, ¹lizhixin@hebeu.edu.cn

收稿日期: 2020年9月29日; 录用日期: 2020年10月13日; 发布日期: 2020年10月20日

摘要

旅游人数是发展旅游业的重要指标, 对河北省未来接待旅游人数的预测一直受到河北省旅游局的重视。本文通过以1990年~2019年河北省旅游数据为依托, 首先, 从市场、景区、政策三个方面分析了河北省旅游业现状, 然后进行了河北省历年来旅游接待人数数据的平稳性和白噪声检验, 分别运用非平稳时间序列的两种残差自回归模型方法(因变量关于时间的回归模型和延迟因变量回归模型)对以往河北省旅游接待人数建立模型。研究发现, 前者模型拟合效果较好, 并对未来旅游人数进行短期预测。最后为促进河北省旅游业的发展提出了一些相关建议。

关键词

时间序列分析, 残差自回归(Auto-Regressive), 旅游接待人数

The Forecast of Tourism Reception Number in Hebei Province Based on Auto-Regressive

Zairan Nie¹, Zhixin Li^{1*}, Zhiguo Li²

¹School of Mathematics and Physics Science and Engineering, Hebei University of Engineering, Handan Hebei

²School of Science, Hebei University of Technology, Tianjin

Email: niezairan@163.com, ¹lizhixin@hebeu.edu.cn

Received: Sep. 29th, 2020; accepted: Oct. 13th, 2020; published: Oct. 20th, 2020

Abstract

The number of tourists is an important index for the development of tourism. The prediction of

*通讯作者。

文章引用: 聂冉冉, 李志新, 李志国. 基于 Auto-Regressive 的河北省旅游接待人数预测研究[J]. 应用数学进展, 2020, 9(10): 1710-1721. DOI: 10.12677/aam.2020.910198

the number of tourists in the future in Hebei province has always been valued by the Hebei provincial tourism administration. Based on the data of tourism in Hebei province from 1990 to 2019, firstly, the present situation of tourism in Hebei province is analyzed from three aspects: market, scenic spot and policy. Secondly, the stability and white noise test of the number of tourism reception in Hebei province over the years are carried out, and two kinds of the residual Auto-Regression models (regression model of dependent variables on time and regression model of delay dependent variables) of the non-stationary time series are used to establish the model of tourism reception in Hebei province in the past. The results show that the former model provides a better fit to the tourism data, which makes a short-term forecast of the number of tourists in the future. Finally, some relevant suggestions are put forward to promote the development of tourism in Hebei province.

Keywords

Time Series Analysis, Auto-Regressive, Number of Tourist Visitors

Copyright © 2020 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

随着时代的进步, 社会和经济的不断发展, 我国人均消费水平急速上升, 在五花八门的消费模式中更多的人青睐于旅游, 如今旅游业已经进入了蓬勃发展时期。旅游业的发展不仅带动了经济的发展, 同时对旅游地区的发展也具有重要的战略性意义。但是随着旅游业规模的扩大, 旅游人数的激增, 人们在外出旅游的同时也遇到了不少阻碍。为了让旅游业得到持续稳定的发展, 必须先搞清楚客源现状和未来的客源发展趋势, 建立科学的、具有现实意义的旅游人数预测模型。而河北省作为我国地势地貌最齐全的省份, 其先天的旅游条件尤为优越, 研究河北省旅游业可以为我国旅游事业的发展提供依据。因此研究河北省旅游接待人数的现状和未来发展状况显得更为重要。

目前, 对旅游人数进行预测是相关部门和研究人员关注的重点。国外方面, Michael D.G [1]利用 Box-Jenkins 方法对夏威夷旅游接待人数进行预测, 来证明该方法作为短期预测方法的有效性。Colladon A F [2]等人对欧洲 7 个城市旅游论坛的相关数据进行了分析, 提出了因子增强的自回归模型来分析旅游相关的大数据, 对预测国际机场抵达人数起到了更大的作用。国内方面常用的预测方法有时间序列预测法、回归分析方法和神经网络, 此外还有灰色预测法等。如孔朝莉[3]利用 2011 年 1 月至 2018 年 8 月共 104 个月的海南旅游人数数据建立 ARIMA(1,1,1) × (1,1,1)₁₂ 模型, 对 2018 年 9 月旅游人数进行预测, 相对误差为 2.06。徐萍丽[4]等在将基于直觉模糊的时间序列模型应用到对杭州宋城主题公园旅游收入的预测上, 此外, 他们还针对提高预测的精准度这一问题对算法进行了改进研究。杨伟芳[5]采用 Regress 函数对海南省 2007 年至 2011 年旅游人数进行回归分析, 该方法具有方便快捷和实用性强等特点。张峰[6]等人构造了具有较好的准确性和鲁棒性的 BP 神经网络模型, 用来预测国内旅游人数。路相林[7]在 BP 神经网络的基础上, 使用经过改良的 BP 神经网络对我国旅游收入总量和旅游入次数进行了预测, 验证了 BP 神经网络模型在研究旅游收入问题上的适用性。程军锋[8]根据 2009 年至 2017 年的陇南市游客数据建立 GM(1,1)模型对未来六年的旅游人数进行了预测, 模型精度较高。余耀峰[9]等人为了弥补传统的统计数据不能实时反映旅游需求变化的不足, 尝试基于 Google 趋势数据对中国入境外国游客的数量进行预测, 并通过将随机森林(RF)模型应用于此任务, 获得了更高的预测精度。

本文研究的主要目的是把时间序列中的残差自回归方法应用到旅游业的人数预测中，通过对历年河北省旅游接待人口总数的分析，找到合适的残差自回归模型，并对河北省未来旅游接待人数的变化趋势进行预测，为相应的管理者提供决策依据，做好下一步管理工作。

2. 河北省旅游业现状

2.1. 市场方面

本文河北省旅游业收入数据均来自 2010 年至 2019 年各地的《国民经济和社会发展统计公报》，具体数据如下表 1 所示。

Table 1. Table of the tourism revenue in Hebei province from 2010 to 2019

表 1. 2010~2019 年河北省旅游业收入表

年份	旅游业收入	年份	旅游业收入
2010 年	914.60 亿元	2015 年	3434.00 亿元
2011 年	1221.30 亿元	2016 年	4654.50 亿元
2012 年	1588.30 亿元	2017 年	6140.90 亿元
2013 年	2010.10 亿元	2018 年	7636.40 亿元
2014 年	2561.50 亿元	2019 年	9313.40 亿元

根据上表数据绘制 2010~2019 年河北省旅游收入折线图：

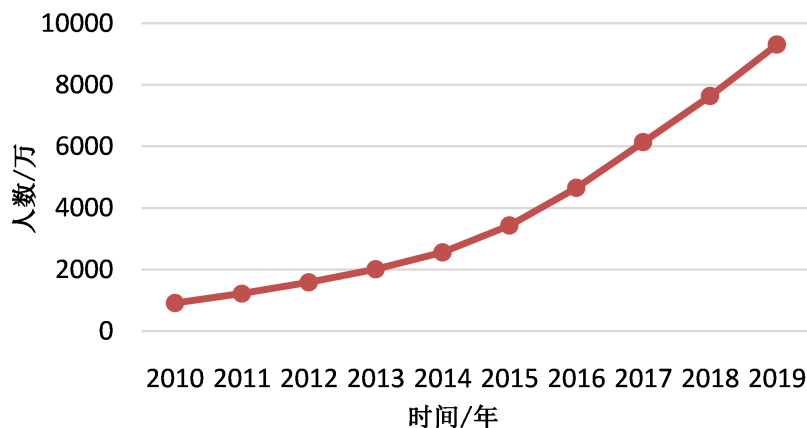


Figure 1. Tourism revenue of Hebei province from 2010 to 2019

图 1. 2010~2019 年河北省旅游业收入

从图 1 可知，2010~2019 年河北省旅游业收入呈现出持续上升的指数增长趋势。自 2014 年起增长速度加快，旅游业收入逐渐成为河北省经济发展中不可或缺的一部分。

2.2. 旅游景区方面

通过 Python 软件，爬取了“去哪儿”网站的共计获取到 136 个 3A 级以上景区的相关数据(景区类型、景区名称、景区 ID、景区等级、景区门票的月销量、景区热度)，本文数据截取时间：2019 年 12 月 25 日。调用“高德地图”的 API 来处理经纬度的换算实现河北省旅游景区的可视化，如下图 2 所示。

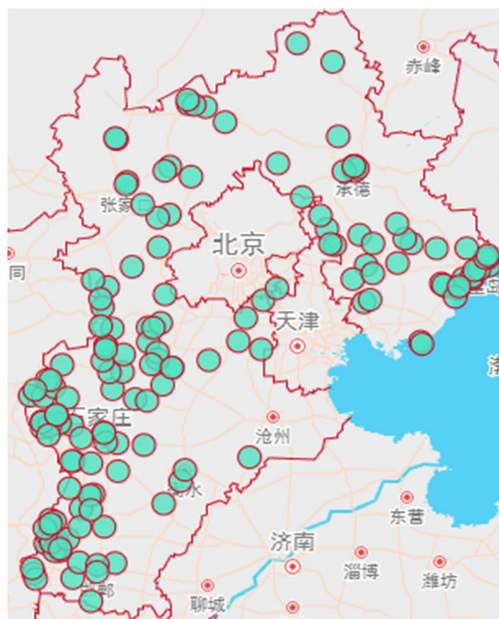


Figure 2. Distribution of 3A and above scenic spots in Hebei province
图 2. 河北省 3A 级以上景区分布

由图 2 可以看出河北省 3A 级以上的旅游景区大多数集中以秦皇岛市为代表的东部沿海地区和保定市 - 石家庄市 - 邯郸市走向的西南部地区。由此可以初步判断石家庄市、秦皇岛市、保定市和邯郸市在河北省内形成了广阔且集中的旅游区域范围，具有绝对旅游资源优势。

根据景区热度数据，本文把收集到的河北省 3A 级以上景区分为 3 个梯度：第 1 梯度为 $0.7 \leq$ 景区热度 < 0.8 、第 2 梯度为 $0.8 \leq$ 景区热度 < 0.9 、第 3 梯度为 $0.9 \leq$ 景区热度 < 1 。并依此绘制河北省 3A 级以上旅游景区热度图，图上颜色越深的区域反映景区的热度就越高，如图 3 所示。



Figure 3. Heat analysis of 3A and above scenic spots in Hebei province
图 3. 河北省 3A 级以上景区热度分析

图3所示景区热度较高的城市有：张家口市、承德市、石家庄市、邯郸市，结合河北省3A级以上旅游景区分布图可知：石家庄市、邯郸市本身旅游景区的基数较大所以热度也相对较高。但是旅游出行受季节温度影响较大，本文收集数据时是正值冬季，那么以清凉避暑为营销目的的东部沿海景区的热度就相对较低，而对于打造冰雪运动休闲区为目标的张家口市和承德市正是吸引游客的最佳季节，所以景区的热度较高。

2.3. 机遇与政策方面

(1) 2022年北京冬季奥运会为旅游业发展提供机遇

第24届冬季奥林匹克运动会将在2022年2月4日至2022年2月20日由中华人民共和国首都北京市和河北省张家口市联合举行，这也是中国历史上首次举办冬季奥运会。京津冀地区则以冬奥会为契机，完成交通路网、基础设施的全面升级，将冰雪旅游资源与体育、娱乐、休闲相结合，打造冰雪旅游产业的新兴地区。

(2) 京津冀协同发展战略——河北省雄安新区

为了疏解北京市的非首都功能、优化京津冀的空间格局和功能定位、构建一体化现代交通网络、扩大环境容量和生态空间、推动公共服务共建共享，在京津冀协同发展战略的大环境下，中共中央、国务院于2017年4月1日决定设立河北省雄安新区。而雄安新区的设立必将促进京津冀三地城际间交通更加便捷、旅游资源更加集中、旅游协作更加密切、旅游市场更加融合，对促进京津冀地区旅游业的协同发展和进一步调整和优化河北省旅游产业的布局，都起着至关重要的作用。

(3) 全面建成小康社会，实现第一个百年奋斗目标

在中国实现小康社会这一伟大构想是早在1979年12月6日由改革开放的总设计师邓小平首次提出，2012年党的十八大报告中首次正式提出全面“建成”小康社会，而2020年则是全面建成小康社会目标实现之年，也是全面打赢脱贫攻坚战收官之年。这些年来在中国共产党的带领下我国紧密结合新的时代条件和人类社会发展的规律出台了众多相关文件大力推进市场经济，同时带动了我国旅游业发展。

3. 方法介绍

残差自回归模型[10] (Auto-Regressive)是研究具有显著的确定性趋势序列时，为了防止残差信息的浪费而提出的一种时间序列拟合模型。本文研究的旅游接待人数是具有明显趋势性的非平稳时间序列模型，因此利用残差自回归模型进行拟合。

Auto-Regressive的基本形式： $x_t = T_t + S_t + \varepsilon_t$ ，其中， T 为趋势效应拟合； S_t 为季节效应拟合； ε_t 为残差序列。首先检验残差序列 $\{\varepsilon_t\}$ 的自相关性，若自相关性不显著(回归模型提取信息充分)，分析结束。若自相关性显著(信息没有提取充分)则需对残差序列 $\{\varepsilon_t\}$ 进行自回归模型拟合，进一步提取相关的信息[11]。拟合的模型如下：

$$\varepsilon_t = \phi_1 \varepsilon_{t-1} + \dots + \phi_p \varepsilon_{t-p} + a_t d \tag{1}$$

其中 a_t 为残差序列 $\{\varepsilon_t\}$ 的残差。若 a_t 为纯随机序列，则模型构造完成，得到的残差自回归模型如下[12]：

$$\begin{aligned} x_t &= T_t + S_t + \varepsilon_t \\ \varepsilon_t &= \phi_1 \varepsilon_{t-1} + \dots + \phi_p \varepsilon_{t-p} + a_t \\ E(a_t) &= 0, Var(a_t) = \sigma^2, Cov(a_t, a_{t-i}) = 0, \forall_i \geq 1 \end{aligned} \tag{2}$$

4. 数据及平稳性白噪声检验

4.1. 数据来源

本文河北省接待旅游人数数据来自河北省统计局官方网站中《河北经济年鉴》(仅有2019年以前数据)，2019年数据来自中商情报网讯。具体数据如下表2所示。

Table 2. Table of the number of tourist receptions in Hebei province from 1990 to 2019
表 2. 1990~2019 年河北省旅游接待人数表

年份	人数	年份	人数	年份	人数	年份	人数	年份	人数
1990 年	1537 万	1996 年	2808 万	2002 年	5985 万	2008 年	9747 万	2014 年	31,368 万
1991 年	2101 万	1997 年	3615 万	2003 年	4477 万	2009 年	12,164 万	2015 年	37,060 万
1992 年	2238 万	1998 年	4016 万	2004 年	7227 万	2010 年	14,851 万	2016 年	46,532 万
1993 年	2412 万	1999 年	4428 万	2005 年	8068 万	2011 年	18,627 万	2017 年	57,234 万
1994 年	2792 万	2000 年	4859 万	2006 年	9053 万	2012 年	22,911 万	2018 年	67,786 万
1995 年	3083 万	2001 年	5316 万	2007 年	10,029 万	2013 年	26,988 万	2019 年	78,188 万

4.2. 平稳性与白噪声检验

本文将河北省旅游接待人数的年度数据序列简记为 X 序列。图 4 是该序列自 1990 年至 2018 年的时序图。该图显示出河北省接待旅客人数呈现出不断上升的趋势，尤其是近几年涨幅速度大大加快。

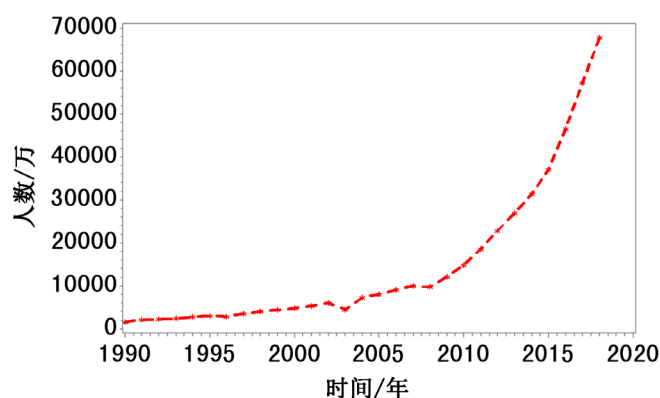


Figure 4. The number of tourist receptions in Hebei province over the years
图 4. 河北省历年旅游接待人数

对 X 序列进行白噪声检验，检验结果如下表 3 所示：

Table 3. Table of white noise test result
表 3. 白噪声检验结果表

延迟阶数	卡方	自由度	Pr > 卡方
6	51.11	6	<0.0001

检验结果显示，在延迟 6 阶情况下 P 值小于 0.0001，即 X 序列为非白噪声序列。

综上所述，可以看出河北省旅游接待人数序列是具有明显的趋势性，但不具备季节性趋势的非平稳非白噪声时间序列。所以该模型实际应为： $x_t = T_t + S_t + \varepsilon_t$ 。其中对于 T_t 的拟合可分为两种方式[13]：

- (1) 自变量为时间 t 的幂函数，称作因变量关于时间的回归模型。
- (2) 自变量为历史观测值，称作延迟因变量回归模型。

下面分别用这两种方式对该序列进行模型拟合。

5. 因变量关于时间的回归模型

由图 4 可以看出, X 序列的趋势并不呈现出线性特征, 而是斜率逐渐增大的曲线。因此, 本文尝试用 $t^2 - t^5$ 拟合模型。经过比较 $t^2 - t^5$ 回归拟合输出统计量的值进行筛选, 本文决定用无趋势项的 t^5 拟合模型。下图 5 为 X 序列的 t^5 模型拟合效果图, 表 4 为 X 与 t^5 的最小二乘法线性回归拟合表。

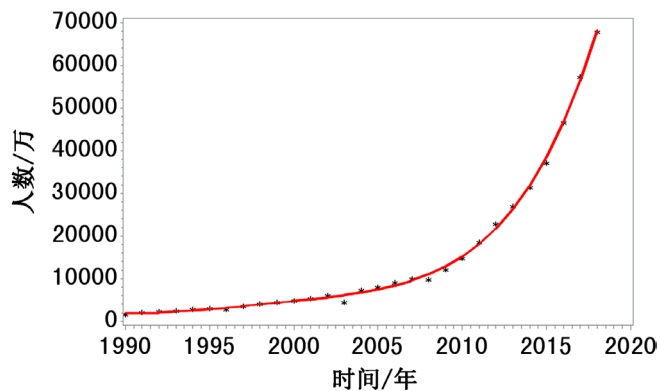


Figure 5. The X sequence t^5 model fitting effect diagram
图 5. X 序列的 t^5 模型拟合效果图

Table 4. Table of X and t^5 least squares linear regression fitting
表 4. X 与 t^5 的最小二乘法线性回归拟合表

统计量	数值
n	28
Durbin-Watson	0.7321
AIC	476.289132
AICC	476.769132
MSE	477.103669
回归 R 方	0.9939
总 R 方	0.9939

从表 4 显示的数据可以发现 R 方的值较高, 说明趋势拟合效果较好, 但是 DW 的值为 0.7321, 证明残差序列 $\{\varepsilon_t\}$ 显著正相关, 由此可知残差序列中信息未提取充分。应对残差序列拟合自相关模型。

由于模型趋势拟合效果较好, 以下仅判定残差部分。如下表 5 和表 6 所示, 残差自相关图显示出拖尾性, 逐步回归消除报告结果显示除了延迟 1 阶的序列值外, 其他延迟阶数均不具有显著的自相关性, 因此延迟 2~9 阶的自相关项被剔除[14]。确定 1 阶残差模型为:

$$\varepsilon_t = \phi_1 \varepsilon_{t-1} + a_t \tag{3}$$

Table 5. Table of residual sequence autocorrelation
表 5. 残差序列自相关表

		自相关估计值	
滞后	协方差	相关	-1, 9, 8, 7, 6, 5, 4, 3, 2, 1, 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 1
0	8,927,566	1.000000	*****

Continued

1	8,459,459	0.947566	*****
2	7,937,452	0.889095	*****
3	7,430,496	0.832309	*****
4	7,039,787	0.788545	*****
5	6,778,762	0.759307	*****
6	6,240,352	0.698998	*****
7	5,630,773	0.630718	*****
8	4,734,864	0.530364	*****
9	3,953,276	0.442817	****

Table 6. Table of stepwise regression elimination report
表 6. 逐步回归消除报告表

自相关估计值				
滞后	估计值	<i>t</i> 值	Pr > <i>t</i>	
9	0.032114	0.08	0.9390	
3	-0.058267	-0.17	0.8684	
4	0.058559	0.23	0.8199	
2	0.125700	0.55	0.5860	
6	0.233031	0.87	0.3950	
7	-0.153035	-0.68	0.5038	
5	-0.258143	-1.94	0.0646	
8	0.111728	1.41	0.1710	

本文选择最大似然估计法来估计具有相关性的趋势参数和残差序列参数，结果如下表 7。

Table 7. Table of maximum likelihood estimation
表 7. 最大似然估计拟合表

统计量	数值
<i>n</i>	28
Durbin-Watson	1.6894
AIC	467.313674
MSE	899568
SSE	23388780.8
回归 R 方	0.9581
总 R 方	0.9977

可以看出整体的 R 方为 0.9977, AIC = 467.313674 说明该模型的拟合度较高。DW 的值等于 1.6894, 查表知 $d_U < 1.6894 < 4 - d_U$, 即残差序列 $\{\varepsilon_t\}$ 已经不具有相关性。参数的估计值及相应的 *t* 检验结果如表 8 所示。

Table 8. Table of parameter estimate (1)

表 8. 参数估计值表(一)

变量	自由度	估计值	标准误差	t 值	Pr > t
t^5	1	0.003284	0.000136	24.19	<0.0001
AR1	1	-0.9339	0.0586	-15.93	<0.0001

通过表 8, 可以发现 t^5 和 AR1 的 P 值均小于 0.0001, 均通过了参数检验。因此可以得到最终模型为:

$$\begin{cases} x_t = 0.003284 \times (t - 1989)^5 + \varepsilon_t \\ \varepsilon_t = 0.9339\varepsilon_{t-1} + a_t \\ E(a_t) = 0, Var(a_t) = 899568, Cov(a_t, a_{t-i}) = 0, \forall_i \geq 1 \end{cases} \quad (4)$$

因变量关于时间的回归模型最终拟合效果如图 6 所示, 其中圆圈为实际值, 红色虚线为拟合值, 上下两条线为置信区间。

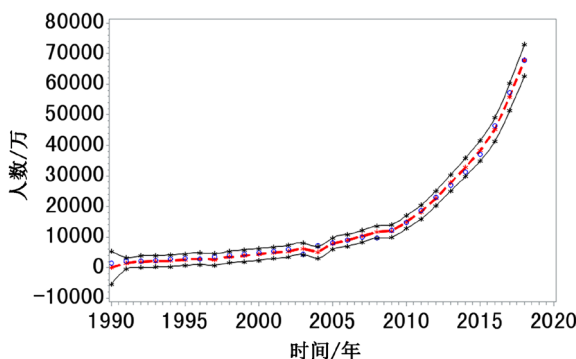


Figure 6. The final model fitting effect

图 6. 最终模型拟合效果

6. 延迟因变量回归模型

首先利用 SAS 中的延迟函数生产序列 X 的 1 阶延迟序列, 然后建立带有延迟因变量的回归模型: $x_t = a + bx_{t-1} + \varepsilon_t$, 并对截距项和延迟一阶系数进行参数估计, 参数估计值及 t 检验结果如下表 9 所示。

Table 9. Table of parameter estimate (2)

表 9. 参数估计值表(二)

变量	自由度	估计值	标准误差	t 值	Pr > t
a	1	-397.501	259.4228	-1.53	0.1374
b	1	1.2140	0.0135	90.02	<0.0001

在显著性水平 0.05 的条件下, 截距项的 P 值为 0.1374 > 0.05, 即不显著, 因此使用逐步回归法继续对模型进行拟合, 拟合统计量的部分结果如下表 10 所示。

Table 10. Table of delay dependent regression analysis

表 10. 延迟因变量回归分析表

统计量	数值
n	28

Continued

Durbin h	-0.1039
Pr < h	0.4586
AIC	469.617822
MSE	1087212
SSE	29354720.8
回归 R 方	0.9980
总 R 方	0.9980

延迟因变量回归分析结果显示总 R 方为 0.9980, AIC = 469.617 822 说明本模型的拟合度较高。由于含延迟因变量, 所以回归结果输出 Durbin h 统计量, 由上表知 Durbin h 统计量的数值已达到 0.4586, 表示残差序列 $\{\varepsilon_t\}$ 显著相关性消除, 即不需要考虑对 $\{\varepsilon_t\}$ 继续拟合自回归模型[15]。最终的参数估计值及相应的 t 检验结果如下表 11。

Table 11. Table of parameter estimate (3)

表 11. 参数估计值表(三)

变量	自由度	估计值	标准误差	t 值	Pr > t
b	1	1.2002	0.0102	117.16	<0.0001

通过表 11, 可以发现延迟 1 阶系数变量 b 的 P 值小于 0.0001, 通过了参数检验。由此可以确定最终拟合模型为:

$$x_t = 1.2002x_{t-1} + \varepsilon_t, \varepsilon_t \sim N(0, 1087212) \tag{5}$$

延迟因变量回归模型最终拟合效果如图 7 所示, 圆圈线为实际值, 红色虚线为拟合值。

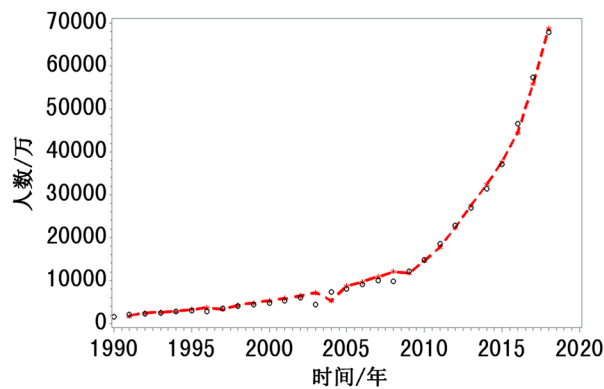


Figure 7. The model fitting effect

图 7. 模型拟合效果

7. 模型比较与选择

模型的比较与选择主要是根据误差平方和最小原则和最小信息量准则, 两个模型的 SSE 值和 AIC 值之前的分析结果已经给出, 现整理到表 12。

根据误差平方和最小原则和最小信息量准则检验结果显示, 因变量关于时间的回归模型的 SSE 值和 AIC 值均小于延迟因变量回归模型, 说明该模型拟合效果更好。从直观的角度考虑因变量关于时间的回

归模型也更容易解释原序列河北省接待旅游人数拥有一个长期的曲线增长趋势的波动规律，所以本文更倾向于选择因变量关于时间的回归模型进行预测。

Table 12. Table of test result
表 12. 检验结果表

模型	SSE	AIC
因变量关于时间的回归模型 $\begin{cases} x_t = 0.003284 \times (t - 1989)^5 + \varepsilon_t \\ \varepsilon_t = 0.9339\varepsilon_{t-1} + a_t \end{cases}$	23388780.8	467.313674
延迟因变量回归模型 $x_t = 1.2075x_{t-1} + \varepsilon_t$	29354720.8	469.617 822

8. 河北省旅游接待人数预测

由图 5 可以看出，拟合值与实际值基本重合，模型的拟合效果良好。通过模型预测得到 2019 年的河北省旅游人数数据为 80,200 万人，与 2019 年的实际值 78188 万人仅仅相差了 2.57%，误差较小。所以认为该模型拟合效果较优异。下面对 2020 年至 2021 年未来两年的河北省旅游人数进行预测。预测结果如下表 13 所示。

Table 13. Table of forecast of the number of tourists in Hebei province from 2020 to 2021
表 13. 2020~2021 年河北省旅游人数预测表

年份	预测结果
2020 年	94,391 万
2021 年	110,541 万

9. 结论

残差自回归模型是一种拟合具有显著的确定性趋势或季节效应的非平稳时间序列方法。它不仅提高了模型的精确度还使结果更切合实际，兼具了时间序列确定性分析和随机性分析的优点。在本文河北省旅游人数的案例中，时间序列呈明显的非线性上升趋势，若仅用一般的时间序列模型进行拟合，残差序列将可能存在自相关性，对信息的提取不够充分，采用残差自回归模型方法大大地降低了拟合误差从而提高了模型的拟合精度。

本文利用残差自回归模型对河北省未来三年的旅游接待人数进行了预测，模型预测效果比较理想。预测结果表明，河北省在未来三年接待的旅游人数将持续上涨，从而河北省旅游业的发展在面临机遇的同时也将承受巨大的压力。旅游人数的增加无疑是河北省经济发展的一个重要增长极。政府和相关旅游管理部门应及时把握机遇大力发展旅游业，为河北省未来的全域旅游发展增加动力的同时也要注意宏观调控，走旅游业的可持续发展道路。

10. 展望

由于 2020 伊始新冠肺炎疫情的爆发给我国国民经济和社会发展造成了不可避免的负面影响，对旅游行业来说更加是一次重创。上述的研究与结论均是在未考虑新冠肺炎疫情影响的前提下得出，接下来的

研究工作将尽可能收集疫情后的旅游业相关数据, 争取更及时、准确的分析出此次疫情对河北省旅游业的具体影响。

基金项目

河北省自然科学基金资助项目(项目编号: A2019402043); 河北省高等学校科学技术研究资助项目(项目编号: QN2019064, ZD2020130)。

参考文献

- [1] Michael, D.G. (1976) Use of the Box-Jenkins Approach to Forecast Tourist Arrivals. *Journal of Travel Research*, **14**, 5-8. <https://doi.org/10.1177/004728757601400402>
- [2] Colladon, A.F., Guardabascio, B. and Innarella, R. (2019) Using Social Network and Semantic Analysis to Analyze Online Travel Forums and Forecast Tourism Demand. *Decision Support Systems*, **123**, 1-11. <https://doi.org/10.1016/j.dss.2019.113075>
- [3] 孔朝莉. 基于 ARIMA 模型的海南月度旅游人数预测[J]. 中国科技信息, 2019(5): 76-78.
- [4] 徐萍丽, 张婷婷, 张鲲. 基于直觉模糊的杭州宋城主题公园旅游收入时序预测的新方法[J]. 经济研究导刊, 2016(7): 159-163.
- [5] 杨伟芳, 田飞. 基于回归模型的海南省旅游人数预测研究[J]. 企业导报, 2012(6): 129-130.
- [6] 张峰, 柳炳祥, 张月. 一种基于 BP 神经网络算法的旅游人数预测方法[J]. 信息与电脑(理论版), 2019, 421(3): 60-62.
- [7] 陆相林. BP 神经网络在中国旅游发展总量预测中的应用[J]. 北京教育学院学报(自然科学版), 2007(1): 8-11.
- [8] 程军锋. 基于灰色理论的陇南市旅游人数预测[J]. 甘肃科技纵横, 2019, 48(4): 59-61.
- [9] Feng, Y.Y., Li, G.W., Sun, X.L. and Li, J.P. (2019) Forecasting the Number of Inbound Tourists with Google Trends. *Procedia Computer Science*, **162**, 628-633. <https://doi.org/10.1016/j.procs.2019.12.032>
- [10] 王燕. 应用时间序列分析[M]. 北京: 中国人民大学出版社, 2000.
- [11] 宋炜晔, 刘志媛, 王飞. 残差自回归模型在旅游经济当中的应用[J]. 现代商贸工业, 2018, 39(6): 23-25.
- [12] 王振龙. 时间序列分析[M]. 北京: 中国统计出版社, 2000.
- [13] 邓美玲, 李小明, 胡荣兴. 国际石油价格之残差自回归模型短期预测[J]. 统计与决策, 2008(22): 146-147.
- [14] 刘松, 李晓妹, 刘健, 刘晓冬, 李向云. Auto-Regressive 模型在全国婴儿死亡率拟合中的应用[J]. 中国卫生统计, 2011, 28(4): 366-368.
- [15] 薛广明, 王洁, 卢若飞. 关于阿里巴巴股票价格的实证研究[J]. 广西民族大学学报(自然科学版), 2015, 21(1): 51-55.