

ARIMA 模型和指数平滑法在我国北方流感样病例预测中的应用和比较

许俐颖¹, 秦国友²

1. 苏州卫生职业技术学院, 江苏 苏州 215009; 2. 复旦大学公共卫生学院, 上海 200032

摘要: **目的** 比较 ARIMA 模型和指数平滑法对我国北方流感样病例的预测效果, 为流感防控提供科学依据。 **方法** 利用我国北方 2012 年第 1 周—2018 年第 17 周的每周流感样病例比例数据拟合建立 ARIMA 模型和指数平滑模型, 预测 2018 年第 18 周—2019 年第 17 周的流感样病例比例, 对预测值与实际值进行比较。 **结果** ARIMA 最优模型为 ARIMA (0,1,1)(2,1,0)_{s2}, 预测的均方根误差 (root mean square error, RMSE)、平均绝对百分比误差 (mean absolute percentage error, MAPE) 和平均绝对误差 (mean absolute error, MAE) 分别为 0.57%、8.98%、0.34%; 指数平滑法的最优模型为简单季节性模型, 预测的均方根误差 (RMSE)、平均绝对百分比误差 (MAPE) 和平均绝对误差 (MAE) 分别为 0.83%、15.24%、0.55%。 **结论** ARIMA (0,1,1)(2,1,0)_{s2} 模型预测精度更高, 可用于我国北方流感样病例的短期预测。

关键词: 流感样病例; ARIMA; 指数平滑; 预测

中图分类号: R511.7 **文献标识码:** A **文章编号:** 1006-3110(2020)07-0889-05 **DOI:** 10.3969/j.issn.1006-3110.2020.07.035

Application and comparison of ARIMA model and exponential smoothing method to the prediction of influenza-like illness cases in northern China

XU Li-ying¹, QIN Guo-you²

1. Suzhou Vocational Health College, Suzhou, Jiangsu 215009, China;

2. School of Public Health, Fudan University, Shanghai 200032, China

Corresponding author: QIN Guo-you, E-mail: gyqin@fudan.edu.cn

Abstract: **Objective** To compare the efficiency of autoregressive integrated moving average (ARIMA) model and exponential smoothing method to predict cases of influenza-like illness (ILI%) in northern China, and to provide a scientific basis for preventing and controlling influenza. **Methods** Data concerning the weekly percentage of outpatient visits for influenza-like illness (ILI) in northern China between the first week of 2012 and the seventeenth week of 2018 were collected to fit and establish ARIMA model and exponential smoothing model. The weekly ILI% from the eighteenth week of 2018 to the seventeenth week of 2019 was predicted, and then the predictive value was compared with the actual value. **Results** The optimal ARIMA model was ARIMA (0,1,1)(2,1,0)_{s2}, and root mean square error (RMSE), mean absolute percentage error (MAPE) and mean absolute error (MAE) of the predictive value were 0.57%, 8.98% and 0.34%, respectively. Simple seasonal model was the optimal model based on exponential smoothing method, and RMSE, MAPE and MAE of the predictive value were 0.83%, 15.24% and 0.55%, respectively. **Conclusions** ARIMA (0,1,1)(2,1,0)_{s2} has a higher forecast precision, and it can be used for short-term prediction of ILI cases in northern China.

Key words: influenza-like illness case; autoregressive integrated moving average; exponential smoothing; prediction

流行性感冒, 是由流感病毒引起的急性呼吸道传染病, 传染性强, 发病率高。曾引起多次世界性大流行, 如何科学的预测流感流行趋势已成为一项重要的

基金项目: 江苏省高职院校教师专业带头人高端研修项目 (2018GRGDYX098)

作者简介: 许俐颖 (1976-), 女, 浙江东阳人, 硕士, 副教授, 研究方向: 公共卫生。

通信作者: 秦国友, E-mail: gyqin@fudan.edu.cn。

课题^[1]。本研究利用我国北方 2012 年第 1 周—2019 年第 17 周的流感样病例监测数据, 分别用自回归和移动平均 (autoregressive integrated moving average, ARIMA) 和指数平滑 (exponential smoothing, ES) 等方法拟合模型, 比较两种模型的预测效果, 确定最优预测模型, 为我国北方流感防控提供科学依据。

1 材料与方法

1.1 资料来源 本研究的资料来源于中国国家流感中心网站上公布的流感监测周报数据(网址: <http://www.chinaivdc.cn/cnic/>;数据路径:国家流感中心-资源中心-流感周报-流感样病例报告)。

1.2 相关定义 流感样病例(influenza-like illness, ILI)的定义是“发热(体温 $\geq 38^{\circ}\text{C}$),伴咳嗽或咽痛之一者”。流感样病例就诊比例(percentage of outpatient visits for influenza-like illness, ILI%)是指哨点医院报告的流感样病例数占哨点医院门急诊病例就诊总数的百分比。

1.3 研究方法 利用我国北方 2012 年第 1 周—2018 年第 17 周的每周 ILI%数据,拟合建立 ARIMA 模型和 ES 模型。并分别用两种模型预测我国北方 2018 年第 18 周—2019 年第 17 周(共 52 周)的 ILI%,与实际值进行比较,评价两种模型的预测效果。

1.4 ARIMA 模型构建^[2-6]

1.4.1 数据的平稳化 绘制时间序列分布图和自相关函数图,分析我国北方 ILI%周次变化资料的随机性、平稳性及季节性。如数据不平稳,则尝试通过差分或数据转换等方式,使序列平稳化。

1.4.2 模型识别和定阶 将模型标记为 ARIMA(p,d,q)(P,D,Q)_s。其中 p 是自回归的阶,d 是差分次数,q 是移动平均的阶。P、D、Q 分别代表季节周期的自回归的阶、差分次数和移动平均的阶,s 指循环长度。根据序列达到平稳时的差分次数确定模型参数 d 和 D,根据差分后序列的自相关函数图和偏自相关函数图确定 p 和 q。季节性参数 P 和 Q 则通过从低阶向高阶逐步试探的方法获得。根据贝叶斯信息准则(Bayesian Information Criterion, BIC),BIC 值小的模型拟合较优。

1.4.3 模型参数估计及诊断检验 对提出的模型进行参数估计和模型诊断。对残差序列进行白噪声检验,计算 Ljung-Box Q 统计量,判断是否为随机性误差。若残差序列为白噪声序列,则认为该模型可以用于预测。

1.5 ES 模型构建^[7] 对我国北方 2012 年第 1 周—2018 年第 17 周的 ILI%数据,用 SPSS 18.0 统计软件逐一构建季节性指数平滑模型,包括简单季节性模型、Holt-Winters 加法模型和 Holt-Winters 乘法模型,并依据决定系数(R^2)、均方根误差(root mean square error, RMSE)、BIC 等指标筛选拟合最优模型。

1.6 统计学分析 用 Excel 2007 建立数据库,用 SPSS 18.0 统计软件的预测模块实现模型构建、诊断和预测。

2 结果

2.1 序列一般情况 将我国北方 2012 年第 1 周—2019 年第 17 周的 ILI%资料绘制成时间序列图,见图 1。其中横坐标的 1 代表 2012 年第 1 周,依次类推,年份变化用参考线标识。根据监测数据显示,我国北方省份的 ILI%在 2013 年出现短暂的上升。序列不平稳有波动,存在明显的季节性,季节性周期约为 52 周,高峰主要集中在冬春季。

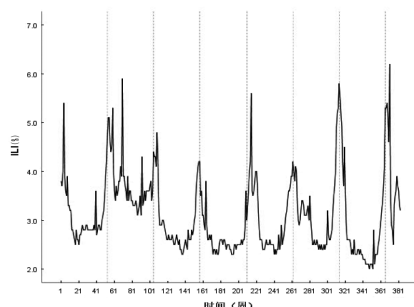


图 1 我国北方 2012 年第 1 周—2019 年第 17 周 ILI%时间序列分布图

2.2 ARIMA 模型构建和优化

2.2.1 序列平稳化处理 对原始数据进行一阶周期长度为 52 周的季节性差分,自相关函数图提示季节性差分后序列不平稳。再进行一阶非季节性差分,差分后序列图曲线在 0 值上下对称波动,自相关系数 1 阶截尾,认为序列平稳。

2.2.2 模型识别与定阶 根据原始数据做 1 阶 52 步差分运算后平稳,考虑构建 ARIMA(p,d,q)(P,D,Q)_s模型进行分析和预测。模型的差分阶数 d=1, D=1, 季节周期 s=52。根据差分后序列的 ACF 图和 PACF 图,自相关系数 1 阶截尾,偏相关系数不截尾,确定 q=1, p=0。根据文献^[5],模型参数中的 P、Q 阶数一般在 2 以内,分别取 0、1、2 从低阶向高阶一一尝试,选择最佳模型。

2.2.3 模型的参数估计和诊断检验 根据以上分析,采取多个模型逐个尝试,同时结合模型统计学检验结果,确定备选模型为 ARIMA(0,1,1)(2,1,0)₅₂, ARIMA(0,1,1)(2,1,1)₅₂, ARIMA(0,1,1)(0,1,1)₅₂。Ljung-Box Q 检验提示三个模型的残差序列均为白噪声序列。模型拟合优度的比较采用决定系数(R^2)、均方根误差(RMSE)和 BIC 值,见表 1。备选模型的参数结果见表 2。ARIMA(0,1,1)(2,1,0)₅₂模型和 ARIMA(0,1,1)(0,1,1)₅₂模型所有参数都有统计学意义,而 ARIMA(0,1,1)(2,1,1)₅₂模型中季节性参数 Q 无统计学意义($P>0.05$)。经综合判断,选择 ARIMA(0,1,1)(2,1,0)₅₂作为最优模型,用于预测。

表 1 备选 ARIMA 模型的拟合优度比较和残差检验

模型	拟合度			Ljung-Box <i>Q</i> 检验		
	<i>R</i> ²	RMSE	BIC	<i>q</i> 值	DF	<i>P</i> 值
ARIMA(0,1,1)(2,1,0) ₅₂	0.711	0.410	-1.680	17.163	15	0.309
ARIMA(0,1,1)(2,1,1) ₅₂	0.711	0.411	-1.657	16.896	14	0.262
ARIMA(0,1,1)(0,1,1) ₅₂	0.699	0.418	-1.663	14.911	16	0.531

表 2 备选 ARIMA 模型的参数估计

参数	ARIMA(0,1,1)(2,1,0) ₅₂			ARIMA(0,1,1)(2,1,1) ₅₂			ARIMA(0,1,1)(0,1,1) ₅₂		
	估计值	<i>t</i> 值	<i>P</i> 值	估计值	<i>t</i> 值	<i>P</i> 值	估计值	<i>t</i> 值	<i>P</i> 值
AR	-	-	-	-	-	-	-	-	-
MA1	0.324	5.586	<0.001	0.329	5.670	<0.001	0.385	6.873	<0.001
SAR1	-0.719	-11.778	<0.001	-0.656	-5.353	<0.001	0.916	-	-
SAR2	-0.550	-9.496	<0.001	-0.522	-6.766	<0.001	-	-	-
SMA	-	-	-	0.091	0.606	0.545	-	2.862	0.005
BIC	-1.680	-	-	-1.657	-	-	-1.663	-	-

2.3 指数平滑模型构建 原始序列呈季节性波动,利用 2012 年第 1 周—2018 年第 17 周的 ILI% 数据,用 SPSS 18.0 构建季节性指数平滑模型,各模型的拟合优度和平滑参数见表 3。简单季节性模型的 RMSE 和 BIC 最小,故认为简单季节性模型是指数平滑模型中的最优模型。该模型的 Ljung - Box *Q* 统计量为 15.643,*P* 值为 0.478,残差为白噪声。模型参数有统计学意义(*t* = 11.571,*P* < 0.001),可用于预测。

表 4 2018 年第 18 周—2019 年第 17 周 ILI% 两种模型预测值与实际值比较

年/周	实际值	ARIMA(0,1,1)(2,1,0) ₅₂			简单季节性指数平滑模型		
	(%)	预测值 (%)	绝对误差	绝对百分比误差 (%)	预测值 (%)	绝对误差	绝对百分比误差 (%)
18/18	2.6	2.3	-0.3	11.5	2.2	-0.4	15.4
18/19	2.3	2.2	-0.1	4.3	2.1	-0.2	8.7
18/20	2.3	2.3	0.0	0.0	2.2	-0.1	4.3
18/21	2.3	2.2	-0.1	4.3	2.1	-0.2	8.7
18/22	2.4	2.3	-0.1	4.2	2.1	-0.3	12.5
18/23	2.4	2.3	-0.1	4.2	2.1	-0.3	12.5
18/24	2.3	2.2	-0.1	4.3	2.1	-0.2	8.7
18/25	2.4	2.3	-0.1	4.2	2.1	-0.3	12.5
18/26	2.3	2.2	-0.1	4.3	2.1	-0.2	8.7
18/27	2.3	2.3	0.0	0.0	2.1	-0.2	8.7
18/28	2.2	2.3	0.1	4.5	2.1	-0.1	4.5
18/29	2.2	2.2	0.0	0.0	2.0	-0.2	9.1
18/30	2.2	2.1	-0.1	4.5	2.0	-0.2	9.1
18/31	2.1	2.2	0.1	4.8	2.0	-0.1	4.8
18/32	2.1	2.2	0.1	4.8	1.9	-0.2	9.5
18/33	2.1	2.2	0.1	4.8	2.0	-0.1	4.8
18/34	2.1	2.1	0.0	0.0	1.9	-0.2	9.5
18/35	2.0	2.1	0.1	5.0	1.9	-0.1	5.0
18/36	2.1	2.1	0.0	0.0	2.0	-0.1	4.8
18/37	2.1	2.1	0.0	0.0	2.0	-0.1	4.8
18/38	2.1	2.1	0.0	0.0	1.9	-0.2	9.5
18/39	2.0	2.2	0.2	10.0	2.4	0.4	20.0
18/40	2.8	2.5	-0.3	10.7	2.2	-0.6	21.4
18/41	2.1	2.3	0.2	9.5	2.1	0.0	0.0
18/42	2.3	2.3	0.0	0.0	2.1	-0.2	8.7
18/43	2.3	2.3	0.0	0.0	2.1	-0.2	8.7
18/44	2.3	2.4	0.1	4.3	2.2	-0.1	4.3
18/45	2.5	2.5	0.0	0.0	2.3	-0.2	8.0
18/46	2.6	2.6	0.0	0.0	2.4	-0.2	7.7
18/47	2.8	2.8	0.0	0.0	2.5	-0.3	10.7
18/48	3.0	2.8	-0.2	6.7	2.7	-0.3	10.0
18/49	3.2	3.0	-0.2	6.3	2.9	-0.3	9.4

表 3 三种季节性指数平滑模型的拟合优度和平滑参数比较

模型	拟合优度			平滑参数 ^a		
	<i>R</i> ²	RMSE	BIC	α	β	γ
简单季节性	0.814	0.316	-2.268	0.600	-	-
Winters 加法	0.814	0.317	-2.248	0.603	-	-
Winters 乘法	0.801	0.327	-2.184	0.688	0.312	-

注:a 各模型的平滑参数均有统计学意义。

2.4 模型预测及预测精度比较 分别用 ARIMA(0,1,1)(2,1,0)₅₂ 模型和简单季节性指数平滑模型对我国北方省份 2018 年第 18 周—2019 年第 17 周共 52 周的 ILI% 进行预测。结果显示,两种方法的 52 周 ILI% 实际值均落在预测值的 95% 置信区间范围内。ARIMA(0,1,1)(2,1,0)₅₂ 模型预测的平均绝对误差 (mean absolute error, MAE) 为 0.34%, 平均绝对百分比误差 (mean absolute percentage error, MAPE) 为 8.98%, 均方根误差 (root mean square error, RMSE) 为 0.57%。简单季节性指数平滑模型预测的平均绝对误差 (MAE) 为 0.55%, 平均绝对百分比误差 (MAPE) 为 15.24%, 均方根误差 (RMSE) 为 0.83%。ARIMA 模型预测结果的各项误差指标均较小, 预测精度优于指数平滑模型 (表 4)。两种模型的预测曲线与实际发病曲线趋势一致, 但 ARIMA 模型的预测曲线有峰值, 更接近有季节性高峰的流感流行规律 (图 2)。

续表 4

年/周	实际值	ARIMA(0,1,1)(2,1,0) _{s2}			简单季节性指数平滑模型		
	(%)	预测值(%)	绝对误差	绝对百分比误差(%)	预测值(%)	绝对误差	绝对百分比误差(%)
18/50	3.4	3.2	-0.2	5.9	3.1	-0.3	8.8
18/51	3.7	3.5	-0.2	5.4	3.3	-0.4	10.8
18/52	4.1	3.9	-0.2	4.9	3.6	-0.5	12.2
19/01	5.3	3.8	-1.5	28.3	3.6	-1.7	32.1
19/02	5.3	3.8	-1.5	28.3	3.6	-1.7	32.1
19/03	5.4	3.9	-1.5	27.8	3.6	-1.8	33.3
19/04	5.1	4.1	-1.0	19.6	3.5	-1.6	31.4
19/05	4.6	4.0	-0.6	13.0	3.3	-1.3	28.3
19/06	6.2	4.5	-1.7	27.4	3.2	-3.0	48.4
19/07	3.4	3.6	0.2	5.9	3.0	-0.4	11.8
19/08	2.9	3.3	0.4	13.8	2.6	-0.3	10.3
19/09	2.8	3.2	0.4	14.3	2.4	-0.4	14.3
19/10	2.5	3.2	0.7	28.0	2.5	0.0	0.0
19/11	3.3	3.4	0.1	3.0	2.5	-0.8	24.2
19/12	3.5	3.4	-0.1	2.9	2.5	-1.0	28.6
19/13	3.6	3.2	-0.4	11.1	2.4	-1.2	33.3
19/14	3.9	2.9	-1.0	25.6	2.3	-1.6	41.0
19/15	3.7	2.6	-1.1	29.7	2.2	-1.5	40.5
19/16	3.6	2.5	-1.1	30.6	2.1	-1.5	41.7
19/17	3.3	2.5	-0.8	24.2	2.5	-0.8	24.2

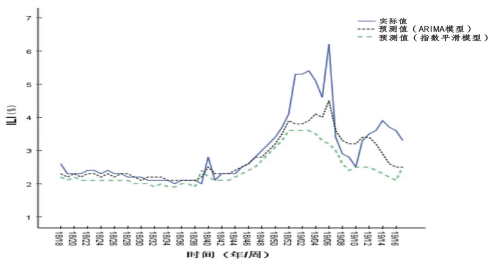


图2 2018 年第 18 周—2019 年第 17 周
ILI%两种模型预测值与实际值比较

3 讨 论

时间序列分析是以预测对象本身的历史数据随时间发展的变化规律建立数学模型外推的预测方法,近年来广泛用于公共卫生和传染病相关的医学研究领域^[4,8]。指数平滑法对整个序列从远到近给予呈指数形式逐渐减弱的权重(即平滑系数),兼容了全期平均和移动平均所长,是国内外广泛使用的一种短期预测方法^[1,7,9]。自回归移动平均(auto regression moving average,ARMA)模型,由自回归模型(AR 模型)与移动平均模型(MA 模型)为基础“混合”构成,是目前研究平稳性时间序列的重要方法。对于非平稳时间序列,ARIMA 模型利用差分将非平稳的时间序列转化为平稳的时间序列,从而将 ARIMA 过程转化为 ARMA 过程。ARIMA 模型被广泛应用于医学研究领域,国内外均有报道^[8]。

本研究分别用 ARIMA 和指数平滑法对我国北方的流感样病例进行预测。预测结果显示,两种模型的预测值与实际值均较接近,拟合曲线与实际发病曲线趋势一致,表明两种模型的预测效果均较好。但相比于简单季节性指数平滑模型,ARIMA (0,1,1)(2,1,0)_{s2}模型的预测结果 RMSE、MAPE、MAE 等指标

均明显较低,认为 ARIMA 模型的预测精度高于指数平滑模型。原因可能在于,ARIMA 模型在建模时综合考虑了序列的趋势、周期变化及随机干扰等情况,并用模型参数进行量化,在数据波动较为明显时,拟合的曲线更能捕捉原始数据的趋势。而指数平滑模型的前提是趋势外延性,即未来一定时期内事物在数量上的演变特征不脱离过去的发展趋势,对数据的转折点缺乏鉴别能力。当流感样病例的发生保持了较长时间的稳定后,突然出现上升或下降趋势时,指数平滑法就难以适应^[10]。相比于简单季节性指数平滑模型,ARIMA (0,1,1)(2,1,0)_{s2}模型更适合用于我国北方流感样病例的发病预测。

虽然 ARIMA 模型的预测精度高于指数平滑模型,但必须指出,ARIMA 模型建模过程也比较复杂。模型的参数常不容易确定,并且对序列的平稳性要求高,在拟合模型之前,需要对非平稳序列进行差分等数据转换。相比于 ARIMA 模型,指数平滑模型的建模过程则更简单一些,不需要太多的用户输入,更容易进行自动化预测^[10]。

有研究^[5,11]指出,应用 ARIMA 模型进行传染病发病的长期预测比较困难,仅适合用于短期预测。但也有研究认为^[12],ARIMA 模型应用于稍长时期的预测(不超过建模时据点的 1/5),其预测效果仍然较好,预测效果稳定。本研究应用 ARIMA 模型进行了 52 周的预测,根据预测结果显示,预测 36 周后,绝对百分比误差明显增大,认为该模型适合用于短期预测。本研究选用 2012 年以后的监测资料作为基础数据进行分析,主要是考虑到在 2012 年国家卫生和计划生育委员会进一步加强了流感监测和报告工作,流感监测网络改善,监测方法统一,监测数据较为稳定^[3,10,13]。本研究