

何延治. 基于时间序列分析的吉林省粮食产量预测模型[J]. 江苏农业科学,2014,42(10):478-479.

基于时间序列分析的吉林省粮食产量预测模型

何延治

(延边大学理学院,吉林延吉 133000)

摘要:吉林省是我国重要的商品粮生产基地,时间序列分析中的 ARIMA(p,d,q)模型可以很好地反映吉林省的粮食生产状况。利用吉林省 1978—2012 年的粮食产量数据,根据时间序列分析法,通过 Eviews 软件得到吉林省粮食产量预测模型,由该模型预测吉林省未来粮食产量呈现稳中增长的势头,但增产幅度逐渐减弱。

关键词:吉林省;时间序列分析;粮食;产量;预测模型

中图分类号: F326.11 **文献标志码:** A **文章编号:** 1002-1302(2014)10-0478-02

吉林省是一个农业大省,是我国重要的粮食主产区,吉林省粮食生产对全国粮食安全具有举足轻重的作用。能够准确预测吉林省的粮食产量对吉林省乃至全国农业发展具有重要的意义^[1]。本研究利用吉林省 1978—2012 年的粮食产量数据,选用时间序列分析中的 ARIMA(p,d,q)模型,建立吉林省粮食产量预测模型,并用该模型拟和吉林省 2011 年和 2012 年的粮食产量,同时预测吉林省 2013 年和 2014 年的粮食产量,这可为指导吉林省粮食生产提供科学的理论依据。

1 时间序列分析方法

时间序列分析方法是由美国著名统计学家博克斯和英国的詹金斯于 1976 年提出,简称 B-J 方法,是一种应用广泛的数量分析方法,主要用于描述和探索现象随时间发展变化的数量规律,是一种时间序列短期预测方法。时间序列预测是通过对预测目标自身时间序列的处理来研究其变化趋势,即通过时间序列的历史数据,揭示现象随时间变化的规律,将这种规律延伸到未来,从而对该现象的未来作出预测。时间序列分析中的预测模型不同于其他计量经济预测模型的一个重要特点是不考虑被研究变量以外的任何其他变量,而是依靠被研究变量本身的外推机制来描述经济变量的变化,强调让数据自己说话,重点分析时间序列本身的随机性^[2-3]。这也是本研究采用该方法的主要原因。

2 时间序列模型

时间序列基本模型有 3 种:自回归模型 (autoregressive model),简称 AR 模型;移动平均模型 (moving average model),简称 MA 模型;自回归移动平均模型 (autoregressive moving average model),简称 ARMA 模型,前 2 种模型是它的特殊形式。这 3 种模型均只适用于平稳的时间序列。但是,实际问题中,许多时间序列并不近似为平稳的时间序列,因此不能直接用非平稳的时间序列建立模型,需要通过差分处理,产生一个平稳的新的时间序列,再用 ARMA 模型建模,这个模型

就是 ARIMA(p,d,q)模型^[4-8]。ARIMA(p,d,q)模型表示为: $\Phi(L)D^d y_t = \theta_0 + \Theta(L)u_t$,其中, $\Phi(L)$ 和 $\Theta(L)$ 分别是 p 阶自回归和 q 阶移动平均算子,它们的根都在单位圆之外, θ_0 是位移项, u_t 是白噪声过程, $D^d y_t$ 表示对非平稳时间序列 y_t 进行 d 次差分后的平稳时间序列。

3 吉林省粮食产量预测模型的建立与预测

3.1 数据收集和整理

选取吉林省 1978—2012 年的粮食产量数据为样本(表 1)。

表 1 吉林省 1978—2012 年的粮食产量

年份	产量 (万 t)	年份	产量 (万 t)	年份	产量 (万 t)
1978	914.70	1990	2 046.52	2002	2 214.80
1979	903.34	1991	1 898.87	2003	2 259.60
1980	859.60	1992	1 840.30	2004	2 510.00
1981	921.91	1993	1 900.90	2005	2 581.21
1982	1 000.04	1994	2 015.70	2006	2 720.00
1983	1 477.98	1995	1 992.40	2007	2 454.00
1984	1 634.46	1996	2 326.60	2008	2 840.00
1985	1 225.26	1997	1 808.30	2009	2 460.00
1986	1 397.71	1998	2 506.00	2010	2 842.50
1987	1 675.81	1999	2 305.60	2011	3 171.00
1988	1 693.25	2000	1 638.00	2012	3 343.50
1989	1 351.29	2001	1 953.40		

为消除数据中可能存在的异方差,利用 Eviews 软件对样本数据取对数生成时间序列 y_t ,再利用 Eviews 软件绘出时间序列 y_t 的相关图,由 y_t 相关图初步看出该时间序列为 1 个非平稳时间序列。对 y_t 进行一次差分运算得时间序列 Dy_t ,利用 Eviews 软件对 y_t 、 Dy_t 进行 ADF 检验^[9],由表 2 可知, y_t 为 1 个非平稳的时间序列, Dy_t 为 1 个平稳的时间序列。这说明吉林省粮食产量适用于 ARIMA($p,1,q$)模型,经过处理而得到的稳定时间序列可运用 ARIMA 模型进行分析。

表 2 y_t 、 Dy_t 的 ADF 检验结果

变量	ADF 检验值	1% 临界值	概率值	结论
y_t	-1.738 0	-3.654	0.403	非平稳
Dy_t	-6.9613	-3.654	0	平稳

收稿日期:2013-12-14

基金项目:国家自然科学基金(编号:31160103)。

作者简介:何延治(1966—),男,吉林龙井人,硕士,副教授,从事计量经济学研究。E-mail:yzhe@ybu.edu.cn。

3.2 模型的建立

ARMA 模型的识别和定阶可通过时间序列的自相关函数值与偏相关函数值获得。AR(p) 模型相关函数值表现为拖尾衰减特征,偏相关函数值在 p 期后出现截止特征;MA(q) 模型自相关函数值在 q 期后出现截止特征,偏相关函数值表现为拖尾衰减特征;ARMA(p, q) 模型的自相关函数值和偏相关函数值都呈拖尾衰减特征,当自相关函数值和偏相关函数值都表现为缓慢的衰减过程时,说明是 1 个 p, q 阶数都大于或等于 1 的 ARMA 模型过程,当自相关函数值出现 q 个峰值后呈指数或正弦衰减,而偏相关函数值出现 p 个峰值后呈指数或正弦衰减,说明是 1 个 ARMA(p, q) 模型^[10]。

利用 Eviews 软件计算出时间序列 Dy_t 的自相关函数值和偏相关函数值(表 3)。由表 3 可知, Dy_t 自相关函数值和偏相关函数值都呈拖尾衰减特征,且均出现 1 个峰值后呈指数或正弦衰减,是 1 个 ARMA(1,1) 模型,从而进一步可确定吉林省粮食产量预测模型为 ARIMA(1,1,1) 模型。

表 3 时间序列 Dy_t 的自相关函数值和偏相关函数值

滞后期	自相关函数值	偏相关函数值	Q 统计量值	概率值
1	0.735	0.735	20.574	0.000
2	0.419	0.095	37.132	0.000
3	0.403	0.034	43.010	0.000
4	0.344	-0.080	47.946	0.000
5	0.310	-0.105	49.846	0.000
6	0.191	0.138	51.479	0.000
7	0.180	0.066	52.985	0.000
8	0.118	-0.082	53.658	0.000
9	0.030	-0.153	53.703	0.000
10	-0.007	-0.015	53.706	0.000

利用 Eviews 软件对时间序列模型 ARIMA(1,1,1) 进行估计,结果见表 4。模型的估计值均通过了 t 检验,具有统计显著性;可决系数不高,这是由于变量差分后损失了很多信息;自回归部分特征根和移动平均部分特征根的倒数都小于 1,这说明该模型具有平稳性和可逆性;模型为: $Dy_t - 0.032 = 0.326(Dy_{t-1} - 0.032) + u_t - 0.985u_{t-1}$, 其中, u_t 是白噪声过程, Dy_t 是时间序列 y_t 的 1 阶差分,以残差 e_{t-1} 代替 u_{t-1} , 虽然 u_t 未知,但 $E(u_t) = 0$, 可取 $u_t = 0$, 得到时间序列 y_t 的 ARIMA(1,1,1) 模型为: $Dy_t = 0.022 + 0.326Dy_{t-1} - 0.985e_{t-1}$, 即 $y_t = 0.022 + 1.326y_{t-1} - 1.326y_{t-2} - 0.985e_{t-1}$, 其中 e_{t-1} 为 $t-1$ 期残差。对模型残差项进行白噪声过程检验,由表 5 可见,残差项的 ADF 检验统计量值为 -5.422 3, 小于 1%、5% 显著水平的临界值,确信残差接近白噪声序列。这说明 ARIMA(1,1,1) 模型为科学合理的预测模型。

表 4 ARIMA(1,1,1) 模型估计结果

变量	系数	标准误差	t	概率值
c	0.032	0.004	8.139	0.000
AR(1)	0.326	0.171	1.998	0.049
MA(1)	-0.985	0.038	-25.642	0.000

注:可决系数 = 0.391,修正后可决系数 = 0.331,回归标准差 = 0.146,对数极大似然估计 = 18.38,因变量均值 = 0.039,因变量方差 = 0.056,残差平方和 = 0.634, DW 统计量 = 2.043,自回归部分特征根倒数 = 0.33,移动平均部分特征根倒数 = 0.98。

表 5 e_t 的 ADF 检验结果

检验水平	t 值	概率值
1% 临界值	-2.9613	0.26
5% 临界值	-1.6725	0.31
ADF 检验值	-5.4223	0

3.3 模型的拟合及预测

利用 ARIMA(1,1,1) 模型,根据表 1 数据,通过 Eviews 软件得到吉林省 2011 年和 2012 年粮食产量分别为 3 214.68 万、3 405.94 万 t,与实际值平均偏差 1.6%,较好地拟合了实际值;同样,通过 Eviews 软件对吉林省 2013 年和 2014 年粮食产量进行预测,结果为 3 462.17 万、3 507.25 万 t,由 ARIMA(1,1,1) 模型预测未来 2 年吉林省粮食产量增产幅度平均为 2.4%,低于过去 5 年 4.7% 的平均增长幅度。这说明吉林省未来粮食产量将保持稳中增长的势头,但按目前生产状况,增产幅度将逐渐减弱。

4 结论

时间序列分析法是一种重要的预测方法,其模型比较简单,对资料的要求比较单一,只需变量本身的历史数据,在实际应用中有广泛的适用性。在应用中需根据所要解决的问题及其特点等因素,综合考虑并选择相对最优的模型。由研究建立的 ARIMA(1,1,1) 模型预测分析可知,吉林省未来粮食产量呈现稳中增长的势头,但增产幅度将逐渐减弱。

为更好地进行吉林省粮食生产,提高吉林省粮食生产的产量,要坚持稳定面积、主攻粮食单产的总方针,向科学技术要产量,要不断加强高标准农田建设,深入实施测土配方施肥,加强病虫害监测与防控能力建设,加快农作物新品种和高效耕作技术的研发推广,促进农产品合理高效流通,充分调动农民的生产积极性。

参考文献:

- [1] 王波,郭夜白,高来斌,等. 最优加权预测在吉林省粮食产量预测中的应用[J]. 吉林农业大学学报,2008,30(5):760-763.
- [2] 潘红宇. 时间序列分析[M]. 北京:对外经济贸易大学出版社,2006:45-97.
- [3] 张树京,齐立心. 时间序列分析简明教程[M]. 北京:北方交通大学出版社,2003:33-84.
- [4] 樊亮,常迎香,李菊梅. 时间序列分析模型在甘肃省 GDP 中的应用[J]. 甘肃科学学报,2009,21(4):140-142.
- [5] 张晓杰,张希良. 时间序列分析模型在山东省粮食总产量预测中的应用[J]. 水土保持研究,2007,14(3):309-311.
- [6] 王延停,杜院录,贾利新. 时间序列分析在粮食产量预测中的应用[J]. 河南科学,2011,29(5):520-523.
- [7] 王燕. 应用时间序列分析[M]. 北京:中国人民大学出版社,2005:45-60.
- [8] 于平福,陆宇明,韦莉萍,等. 基于小波广义回归神经网络的粮食产量预测模型[J]. 湖北农业科学,2011,50(10):2135-2137.
- [9] 张晓峒. 计量经济学软件 EViews 使用指南[M]. 天津:南开大学出版社,2003:232-242.
- [10] 张晓峒. 计量经济学基础[M]. 天津:南开大学出版社,2001:282-341.