

基于 ARMA 模型的河北省“十一五”时期 三次产业就业人数预测分析

张丽峰

(东北大学秦皇岛分校 经济系, 秦皇岛 066004)

摘 要:随着经济的不断发展,三次产业产值结构的调整必然带来就业结构的调整。通过建立河北省三次产业就业人数的时间序列 ARMA 模型,对河北省“十一五”时期三次产业就业人数进行了预测,分析了其特点,并提出了优化调整河北省三次产业就业结构的对策。

关键词:“十一五”时期;三次产业;就业人员;ARMA 模型;河北

中图分类号:F201 **文献标识码:**A **文章编号:**1002-980X(2008)03-0114-05

经济的发展会伴随着产业结构的变化,即人均国民收入水平的提高会使消费结构发生改变,产业结构会由第一产业向第二产业移动,当国民收入水平进一步提高时,产业结构将向第三产业转移,与此同时,劳动力的就业结构也会相应变化。我国改革开放以来,河北省的经济实现了持续、快速的增长,产业结构得到了进一步调整和优化,从三次产业的产出结构看,河北省第一产业增加值比重明显下降,第二产业依然在国民经济中占主导地位,第三产业迅速发展。1978 年,河北省三次产业产出结构为 28.5 50.5 21,2005 年为 14.9 51.8 33.3,自 20 世纪 90 年代以来,第二产业一直在国民经济处于首要产业地位,第三产业在国民经济中的地位也得到了进一步加强,在 20 世纪 90 年代中期超过了农业成为河北省国民经济的第二大产业。

产业结构的调整必然导致就业结构的变动。河北省三次产业就业人员结构由 1978 年的 76.88 13.88 9.24 调整为 2005 年的 45.1 30.2 24.7;第一产业就业人员比重大幅度下降,该比例从 1978 年的 76.88% 下降到 2005 年的 45.1%,下降了 31.78%;第二产业、第三产业的就业人员比重则明显上升,第二产业就业人口比例从 1978 年的 13.88% 上升到 2005 年的 30.2%,上升了 16.32%,第三产业的就业人口比例从 1978 年的 9.24% 上升至 2005 年的 24.7%,上升了 15.46%,就业结构调整逐渐趋向合

理。本文通过建立三次产业就业人数的时间序列 ARMA 模型,对河北省“十一五”时期三次产业的就业人数进行预测,试图找到其变化趋势,旨在为政府制定合适的产业政策和就业政策提供依据。

1 ARMA 模型的建立与分析

ARMA 模型^[1] (Auto-Regressive and Moving Average Model) 也称为博克斯 - 詹金斯法或 B-J 法,是由美国统计学家 Box 和英国统计学家 Jenkins 在 20 世纪 70 年代提出的一种时间序列分析模型,即自回归移动平均模型。ARMA 模型的基本思想是:某些时间序列是依赖于时间 t 的一族随机变量,构成该序列的单个序列虽然具有不确定性,但整个序列的变化却有一定的规律性,可以用相应的数学模型近似描述。通过对该数学模型分析研究,能够更本质地认识时间序列的结构与特征,达到最小方差意义下的最优预测。

1.1 变量和数据的选取

本文对河北省三次产业的就业人员选取第一产业、第二产业、第三产业的就业人员数,单位是万人,时间段是 1978—2005 年。第一产业就业人员序列用 Y_P 表示,第二产业就业人员用 E_P 表示,第三产业就业人员用 S_P 表示。为了消除数据的剧烈波动对其取自然对数,分别表示为 $L Y_P$ 、 $L E_P$ 、 $L S_P$,一阶差分记分别记为 $D L Y_P$ 、 $D L E_P$ 、 $D L S_P$,二阶

收稿日期:2007-10-13

基金项目:河北省教育科学研究“十一五”规划课题(06020645)研究成果之一

作者简介:张丽峰(1969—),女,河北秦皇岛人,东北大学秦皇岛分校经济系讲师,经济学博士,主要研究方向:宏观经济、计量经济、能源经济等。

差分记为 $DDLYP$ 、 $DDLEP$ 、 $DDLSP$ 。本文数据来源于河北省历年经济统计年鉴。

1.2 平稳性检验与模型识别

下面通过对三次产业就业人员序列、一阶差分序列和二阶差分序列各自的相关图和偏自相关图进行分析,来判断它们的平稳性并识别模型的形式^[2,3]。

根据第一产业就业人员序列 LYP 的自相关图可知,其自相关系数没有很快趋于 0,说明该序列是非平稳的^[3]。

从图 1 可以看出,第一产业就业人员序列 $DLYP$ 的自相关系数很快趋于 0,说明该序列是平稳的。由图 1 和图 2 可知, $DLYP$ 序列的偏自相关函数在第 4 期以后很快趋于 0,可以考虑 $AR(4)$,自相关函数具有拖尾性质,因此可初步建立 $AR(4)$ 时间序列模型。

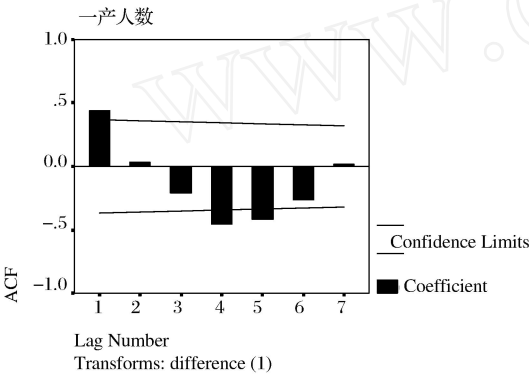


图 1 $DLYP$ 序列的自相关图

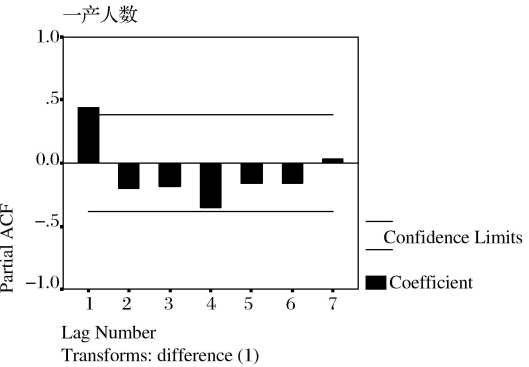


图 2 $DLYP$ 序列的偏自相关图

根据第二产业就业人数序列 LEP 的自相关图可知,其自相关系数没有很快趋于 0,说明该序列是非平稳的^[3]。

从图 3 可以看出,第二产业就业人数序列 $DLEP$ 的自相关系数很快趋于 0,说明该序列是平稳的。由图 3 和图 4 可知, $DLEP$ 序列的偏自相关函数在第 2 期以后很快趋于 0,可以考虑 $AR(1)$,自

相关函数在第 1 期显著不为 0,可以考虑 $MA(1)$,因此可初步建立 $ARMA(1,1)$ 时间序列模型。

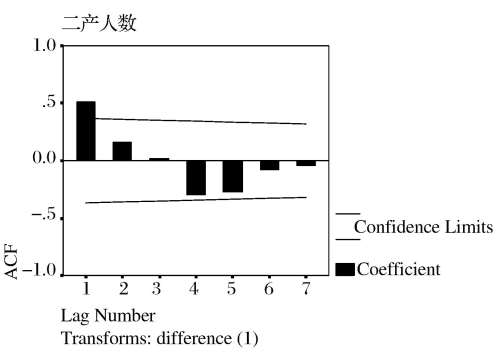


图 3 $DLEP$ 序列的自相关图

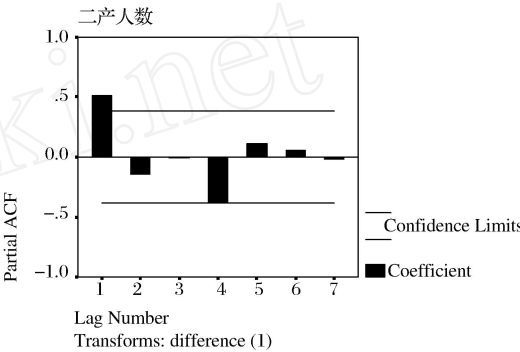


图 4 $DLEP$ 序列的偏自相关图

根据第三产业就业人员序列和一阶差分序列的自相关图可知,第三产业就业人数序列 LSP 和一阶差分序列 $DLS P$ 的自相关系数没有很快趋于 0,说明序列是非平稳的^[3]。

从图 5 可以看出,第三产业就业人数二阶差分序列 $DDLSP$ 的自相关系数很快趋于 0,说明序列是平稳的。由图 5 和图 6 可知, $DDLSP$ 序列的偏自相关函数在第 1 期以后很快趋于 0,可以考虑 $AR(1)$,自相关函数在第 2 期显著不为 0,可以考虑 $MA(2)$,因此可初步建立 $ARMA(1,2)$ 时间序列模型。

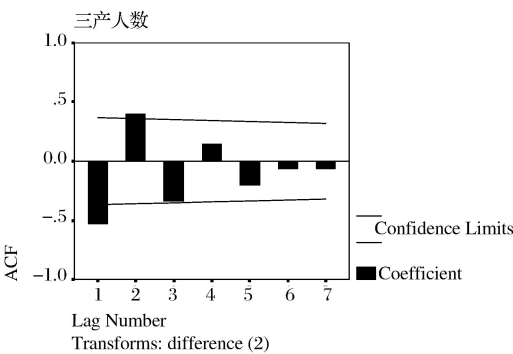


图 5 $DDLSP$ 序列的自相关图

因此,从河北省三次产业的自相关图和偏自相关图可以看出,第一产业就业人员序列一阶差分是

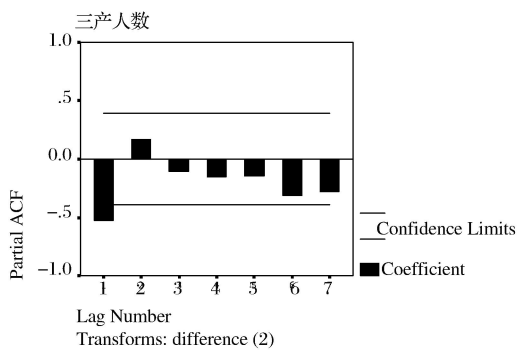


图 6 DDLSP 序列的偏自相关图

平稳的,第二产业就业人员序列一阶差分是平稳的,第三产业就业人员序列二阶差分是平稳的。

1.3 参数检验

为了检验模型的预测效果,将 2003—2005 年这 3 年的数据留出,作为评价预测精度的参照对象,将建模的样本期调整为 1978—2002 年。利用 Eviews5.1 软件^[3],通过 OLS 估计分别得出河北省三次产业就业人员的估计方程。

(1)第一产业就业人员的估计方程 估计方程如下:

$$DLYP = 0.3767AR(1) - 0.1077AR(2) - 0.0459AR(3) - 0.3586AR(4)$$

(1.61) (-0.43) (-0.18) (-1.57) ° (1)

表 1 模型(2)残差的自相关和偏自相关函数表

Autocorrelation			Partial Correlation			Lag Number	AC	PAC	Q-Stat	Prob
.	.	.	.	.	.	1	-0.057	-0.057	0.0746	0.785
.	*	*	.	*	*	2	-0.248	-0.252	1.5827	0.453
.	.	.	.	.	.	3	0.010	-0.023	1.5856	0.663
.	.	*	.	.	*	4	0.226	0.175	2.9953	0.559
.	*	*	.	*	*	5	-0.139	-0.125	3.5615	0.614

由表 1 可知, $Q = 3.5615 < \chi^2_{0.05}$,另外当滞后期小于 5 时,残差序列的自相关系数都落入随机区域,表明残差序列是纯随机的。

(2)第二产业就业人员的估计方程 估计方程如下:

$$DLEP = 0.5562AR(1) + 0.2089MA(1)$$

(2.23) (0.69)

$R^2 = 0.21; Adj R^2 = 0.17; AIC = -2.62;$
 $SC = -2.52; D.W. = 2.03.$

Inverted AR Roots .56
Inverted MA Roots -.21

式(3)中,括号中的数值为 t 值,去掉不显著的

$R^2 = 0.35; Adj R^2 = 0.23; AIC = -3.8;$
 $SC = -3.6; D.W. = 2.0.$

Inverted AR Roots .63 + .59i .63 - .59i
 - .44 - .54i - .44 + .54i

式(1)中,括号中的数值为 t 值,去掉不显著的变量,再进行回归得:

$$DLYP = 0.3165AR(1) - 0.3971AR(4)$$

(1.61) (-2.07)

$R^2 = 0.33; Adj R^2 = 0.3; AIC = -3.98;$
 $SC = -3.88; D.W. = 1.99.$

Inverted AR Roots .65 - .55i .65 + .55i
 - .49 + .55i - .49 - .55i

从模型(2)的检验参数看,各参数均通过显著性检验,模型的调整判别系数在逐渐提高,AIC 值和 SC 值在逐渐减少,D.W. 值表明模型不存在自相关。各滞后多项式的倒数根都在单位圆内,说明过程既是平稳的也是可逆的。

下面对模型(2)的适合性进行检验,即对模型的残差序列进行白噪声检验。若残差序列不是白噪声序列,意味着残差序列中还存在有用的信息没被提取,需要进一步改进模型。表 1 是模型(2)的残差的自相关和偏自相关函数表。

变量,再进行回归得:

$$DLEP = 0.6710AR(1)$$

(4.36) (4)

$R^2 = 0.2; Adj R^2 = 0.2; AIC = -2.69;$
 $SC = -2.65; D.W. = 1.89.$

Inverted AR Roots .67

从模型(4)的检验参数看,各参数均通过显著性检验,模型的调整判别系数在逐渐提高,AIC 值和 SC 值在逐渐减少,D.W. 值表明模型不存在自相关。各滞后多项式的倒数根都在单位圆内,说明过程既是平稳的也是可逆的。

滞后多项式的倒数根。

下面对模型(4)的适合性进行检验,即对模型的残差序列进行白噪声检验。若残差序列不是白噪声序列,意味着残差序列还存在有用的信息没被提取,

需要进一步改进模型。表 2 是模型(4)残差的自相关和偏自相关函数表。

表 2 模型(4)残差的自相关和偏自相关函数表

Autocorrelation	Partial Correlation	Lag Number	AC	PAC	Q - Stat	Prob
. * .	. * .	1	- 0. 112	- 0. 112	0. 3270	0. 567
. * .	. * .	2	- 0. 126	- 0. 141	0. 7642	0. 682
. * * .	. * .	3	0. 212	0. 187	2. 0624	0. 560
* * * .	* * * .	4	- 0. 339	- 0. 333	5. 5392	0. 236
. * .	. * * .	5	- 0. 184	- 0. 219	6. 6222	0. 250
. * .	. * .	6	0. 084	- 0. 089	6. 8636	0. 334

由表 2 可知, $Q = 6. 8636 < \chi^2_{0.05}$ 。另外,当滞后期小于 6 时,残差序列的自相关系数都落入随机区域,自相关系数的绝对值几乎都小于 0. 1,与 0 无显著的差异,表明残差序列是纯随机的。

(3)第三产业就业人员的估计方程 估计方程如下:
 $DDLSP = - 0. 6877AR(1) + 0. 3785MA(1) + 0. 3564MA(2);$ (5)

(- 2. 59) (1. 28) (1. 53)
 $R^2 = 0. 36; Adj R^2 = 0. 3278; AIC = - 3. 07;$
 $SC = - 2. 92; D. W. = 2. 1。$

Inverted AR Roots - . 69
Inverted MA Roots - . 19 + . 57i - . 19 - . 57i
式(5)中,括号中的数值为 t 值,去掉不显著的变量,再进行回归得:

$DDLSP = - 0. 4512AR(1) + 0. 2762MA(2);$ (6)
(- 2. 17) (1. 22)

$R^2 = 0. 39; Adj R^2 = 0. 3286; AIC = - 3. 11;$
 $SC = - 3. 01; D. W. = 1. 91。$

Inverted AR Roots - . 45
去掉不显著的变量,再进行回归得:
 $DDLSP = - 0. 5659 AR(1)。$ (7)
(- 3. 15)

$R^2 = 0. 3211; Adj R^2 = 0. 3211; AIC = - 3. 14;$
 $SC = - 3. 09; D. W. = 1. 7。$

Inverted AR Roots - . 57
从模型(7)的检验参数看,各参数均通过显著性检验,模型的 AIC 值和 SC 值在逐渐减少,D. W. 值表明模型不存在自相关。各滞后多项式的倒数根都在单位圆内,说明过程既是平稳的也是可逆的。

下面对模型(7)的适合性进行检验。表 3 是模型(7)残差的自相关和偏自相关函数表。

表 3 模型(7)残差的自相关和偏自相关函数表

Autocorrelation	Partial Correlation	Lag Number	AC	PAC	Q- Stat	Prob
. * .	. * .	1	0. 145	0. 145	0. 5270	0. 468
. * .	. * .	2	0. 174	0. 157	1. 3302	0. 514
. * * .	. * * .	3	- 0. 235	- 0. 292	2. 8621	0. 413
. * * .	. * * .	4	- 0. 221	- 0. 198	4. 2999	0. 367
* * * .	. * * .	5	- 0. 322	- 0. 202	7. 5194	0. 185

由表 3 可知, $Q = 7. 5194 < \chi^2_{0.05}$ 。另外,当滞后期小于 5 时,残差序列的自相关系数都落入随机区域,表明残差序列是纯随机的。

为了检验模型的预测效果,利用模型(2)、模型(4)和模型(7)对 2003—2005 年河北省的三次产业就业人数进行试预测。结果见表 4。

从表 4 可知,预测值与实际值差别不大。预测精度的情况为:第一产业就业人员的 MAPE 为 2. 69,第二产业为 3. 62,第三产业为 2. 47。一般认

为,当 MAPE 值低于 10 时预测精度较高。
由于模型的合适性通过了检验,试预测精度也很高,因此可判断该模型是可行的,可以用其进行预测。

1. 4 预测结果

利用模型(2)、模型(4)和模型(7)对“十一五”时期(2006—2010 年)河北省的三次产业就业人数进行预测,预测结果见表 5。

表 4 2003—2005 年河北省三次产业就业人员试预测值

万人

年份	第一产业		第二产业		第三产业	
	预测值	实际值	预测值	实际值	预测值	实际值
2003	1 657. 054	1 672. 260	949. 413	942. 840	855. 042	855. 130
2004	1 645. 518	1 612. 850	963. 277	992. 740	866. 109	911. 120
2005	1 642. 576	1 562. 300	972. 694	1 048. 400	877. 805	856. 600

注:预测值为模型的预测结果,实际值来源于河北省历年相关统计年鉴。

表 5 2006—2010 年河北省三次产业就业人员预测值

万人

年份	第一产业	第二产业	第三产业
2006	1 551. 700	1 087. 488	993. 624
2007	1 544. 798	1 114. 530	1 036. 552
2008	1 564. 941	1 133. 051	1 086. 839
2009	1 591. 371	1 145. 652	1 136. 295
2010	1 604. 161	1 154. 185	1 189. 936

2 结果分析

由表 5 可知:“十一五”时期河北省第一产业的就业人数呈波动变化趋势,但其所占的比重呈下降趋势,2006 年为 1551. 7 万人,占三次产业就业人数的 42. 7%,2007 年有所下降,到 2010 年为 1604. 161 万人,占 40. 6%,下降了 2. 1%;第二产业就业人数总量虽呈上升趋势,但其所占的比重呈下降趋势,2006 年为 1087. 488 万人,占三次产业就业人数的 30%,2010 年为 1154. 185 万人,占 29. 4%,下降了 0. 6%;第三产业就业人数总量和所占比重均呈现上升趋势,2006 年为 993. 624 万人,占三次产业就业人数的 28. 9%,2010 年为 1189. 936 万人,占 30. 1%,上升了 0. 2%。总体上来讲,三次产业就业人数的变化趋势大体符合产业结构的变化趋势,但仍存在差距。第一产业就业人数所占的比重逐步下降,但仍达到 40 %左右,第二产业、第三产业的就业人数总量逐年上升,但第二产业、第三产业就业人

数所占的比重仍偏低。

3 建议

根据前文的分析结果并结合河北省经济发展的特点,本文对河北省三次产业就业结构的优化调整提出如下对策建议:

- 1)要在推进科技进步,发展资本、技术型产业的同时,发展劳动密集型产业,大力扶持劳动密集型中小企业的发展,促进“小企业,大就业”。
- 2)要大力发展非公有制经济,特别是私营经济和个体经济,充分发挥非公有制经济在吸纳就业方面的主渠道作用。
- 3)要发展知识密集与劳动密集相结合的高新技术产业,创造新的就业岗位。
- 4)要大力发展第三产业,通过第三产业的发展把第一产业的就业人员解放出来,以促进就业,在保持经济稳定增长的同时,实现人与社会、经济的协调发展。

参考文献

[1] 迪博尔德. 经济预测[M]. 张涛,译. 2 版. 北京:中信出版社,2003:43-50.

[2] 张晓峒. 计量经济分析[M]. 北京:经济科学出版社,2000:136-173.

[3] 高铁梅. 计量经济分析方法与建模[M]. 北京:清华大学出版社,2006:126-137.

Analysis and Forecast on Employment Number of Thrice Industry in Hebei during the Eleventh Five-Year Plan Period Based on ARMA Model

Zhang Lifeng

(Northeastern University at Qinhuangdao ,Qinhuangdao Hebei 066004 ,China)

Abstract : With the rapid economy development ,the change of thrice industrial structure brings about the adjustment of employment structure. Based on ARMA model ,this paper forecasts the employment number of thrice industry in Hebei during the eleventh five-year plan period ,analyzes its characteristics and puts forward some suggestions on adjusting and optimizing industrial structure.

Key words : the eleventh five-year plan period ; thrice industry ; employment number ; ARMA model ; Hebei