

基于 ARIMA 模型对我国就业状况的预测

■马 跃 厦门大学经济学院

[摘要] 本文采用求和自回归移动平均模型(ARIMA)对我国1952年~2006年的就业人数的数据进行时间序列分析,结果显示ARIMA(2,1,1)模型提供了较准确的预测结果,可用于未来的预测,就此可为我国社会保障部门提供一些参考数据。

[关键词] 时间序列分析 就业人员合计 ARIMA模型 未来预测

充分就业与经济增长、物价稳定、国际收支平衡并列为各国政府管理经济的四大目标。近年来,我国的就业压力越来越大,受到社会各界的广泛关注,成为各级政府面临的严峻挑战和亟待解决的主要难题;增加就业的难点,体制转轨释放出的压力,信息化进程加快使就业的难度加大,就业结构性矛盾突出。为了准确预测我国就业人数的发展趋势,使所建模型既要满足实际的要求,也满足统计方法理论的要求,故文章采用自回归移动平均模型来建立我国就业人数的预测模型。

一、RIMA 模型的构建思想

ARIMA 是一类常用的随机时间序列模型,由博克斯、詹金斯创立。它是一种精度较高的时间序列短期预测方法,其基本思想是:某些时间序列是依赖于时间 t 的一簇随机变量,构成该时序的单个序列值虽然具有不确定性,但整个序列的变化却有一定的规律性,可以用相应是数学模型近似描述。通过对该序列的分析研究,能够更本质的认识时间序列的结构与特征,达到最小方差意义下的预测。

ARIMA(p, d, q)模型中,AR指自回归,P为模型的自回归项数;MA为移动平均,q为模型的移动平均项数;d为时间序列成为平稳之前必须取其差分的次数。其一般的表达式为:

$$Y_t = a_0 + a_1 Y_{t-1} + a_2 Y_{t-2} + \dots + a_p Y_{t-p} + b_0 U_t + b_1 U_{t-1} + b_2 U_{t-2} + \dots + b_q U_{t-q}$$

二、ARIMA 模型的应用

根据 ARMA 模型的前提条件,建立模型的时间序列方法是以平稳随机时间序列为前提的。选取我国 1952 年~2006 年就业人数合计的数据,令其为 W ,并对数据进行平稳性检验与处理。

1. 对时间序列 W 进行分析

对其做折线图,观察到曲线向右上方倾斜,并且在其后波动幅度不一致,说明序列存在增长趋势。对数据进行单位根检验,从检验结构($P = 0.9966$)中得出数据没有通过 ADF 检验,因此该序列是非平稳的时间序列。说明我国就业人数的变化受多种因素

影响而不能采用固定模式进行分析预测。

2. 对 W 进行平稳化处理

对 $\{W\}$ 序列进行差分,经检验发现需要对 $\{W\}$ 序列进行一阶差分才能使序列达到平稳,ADF 检验量的 $P=0.0000$ 。

3. 模型识别

即选择用 AR(P),MA(q),还是用 ARMA(p, q)模型对平稳的时间序列模型进行估计。通过前文我们已经知道 $I(d)$ 的阶数为 1,即 $d=1$ 。现对 ARMA 模型进行定阶分析。首先利用 ACF 图和 PACF 图进行模型阶数分析,初步确定为 ARMA(2,1)模型。此模型效果指标如下表:

	AIC	SC
ARMA(2,1)	16.96937	17.04442

从下表的数据可以证实其参差接近白噪声序列:

ACF 和 PACF 图的观察

LAG	1	2	3	4	5	6	7	8	9
AC	-0.209	-0.232	0.099	-0.126	0.095	0.023	-0.068	0.047	0.017
PAC	-0.209	-0.289	-0.028	-0.198	0.095	-0.028	-0.018	0.006	0.036
Q-stat	2.4061	5.4409	6.0010	6.9310	7.4660	7.4989	7.7860	7.9293	7.9493

可拟核出此模型为: $Y_t = 0.483293Y_{t-2} + 0.342998U_{t-1}$

三、ARIMA 的预测与分析

对 $\{W\}$ 作出 2007 年~2010 年的预测值:

2007 年~2010 年就业人数预测值

年份	2007	2008	2009	2010
预测值	76752.49	77030.38	77200.74	77335.04

由于 ARIMA 模型模型的自身存在先天性缺陷:随着预测期的延长,其预测误差也会逐渐增大,但在短期内它的预测还是比较准确的,而且与其他的预测方法相比,其预测的准确程度还是比较高的,尤其在短期预测方面。

希望本文的预测结果能给我国关于就业情况的分析与改善提供一些依据,使我国可以更好地优化就业结构,提供更多的工作岗位,维护社会的稳定,促进我国经济加快发展进程。