International Journal of

Dairy Science

ISSN 1811-9743

Academic Journals Inc.

www.academicjournals.com

International Journal of Dairy Science 5 (3): 204-210, 2010
ISSN 1811-9743 / DOI: 10.3923/ijds.2006.63.69

Unit Root Test and Forecast of Milk Production in Pakistan*

Heman D. Lohano and Fateh M. Soomro
Sindh Agriculture University, TandoJam, Pakistan

Abstract: This study forecasts milk production in Pakistan using historical time series data. Random walk model with drift and trend-stationary autoregressive model are considered for estimation. Unit root test is a diagnostic test for identifying whether a time series is represented as a random walk model with drift or a trend-stationary autoregressive model. Results showed that milk production in Pakistan is characterized as a random walk model with drift, which indicates that the shocks to production in a year have permanent effect on the level of future production. The forecast results showed that the annual milk production is expected to grow at 4.17% per annum.

Key words: Unit root, random walk, forecast, autoregressive model, milk production

Introduction

Livestock is an important subsector of agriculture in Pakistan. In 2004-05, the share of livestock was 46.8% in agriculture value added and 10.8% in the Gross Domestic Product (Anonymous, 2005). In GDP accounting, livestock includes the value of milk, draught power, dung, urine, wool, hair, net sale of animals for slaughtering, natural growth of animals and poultry products. Among these products, milk is a major product contributing about 52% in the value of livestock subsector (Anonymous, 2004). Raising livestock for milk and other products is a key economic activity in Pakistan as it provides income support to most farmers, a portfolio for diversification in farming, employment to landless farmers and it is a major source of income in rural areas predominantly in arid areas.

In the world, Pakistan has become the forth largest milk producer after India, USA and Russian Federation (Anonymous, 2005b). In 2004-05, milk production in Pakistan was 29.472 million tonnes with per capita production of 191 kg (Anonymous, 2005a). There is, however, still rising demand for milk especially in big cities including Karachi, Hyderabad, Peshawar and Quetta, as reflected from very high price of milk in these cities compared to rural areas. Furthermore, milk has a high income elasticity of demand for lower income classes (Anonymous, 2001). Thus, there will be more demand for milk with population growth, economic growth, urbanization and awareness of nutritional requirements.

As milk production has significant importance on both supply and demand sides, Government needs accurate and advance information about the future status of milk production for short and long term planning. Forecasting of milk production may assist the policy makers in formulating policies regarding support services and market regulations.

Milk production in Pakistan has grown from 7.800 million tonnes in 1971-72 to 29.472 million tonnes in 2004-05 (Anonymous, 2005a). Trend in time series may arise from the positive drift term of a random walk model or from a deterministic time trend of a trend-stationary autoregressive model.

Corresponding Author: Dr. Heman D. Lohano, Assistant Professor, Department of Agricultural Economics, Sindh Agriculture University, TandoJam, Pakistan Tel: 92-300-304-2034

Originally Published in International Journal of Dairy Science, 2006

In a random walk model with drift, the shocks to production have permanent effect for the level of future production. In a trend-stationary process, the shocks have temporary effect with the lost production made up during the recovery (Hamilton, 1994).

Milk production is stochastic and its growth over time is subject to various shocks including animal death due to disease or natural disaster, low reproduction, fodder and water availability. In a year, if there is an event of animal deaths or low reproduction, then this will have a permanent effect on future production. This is because these shocks have reduced the animal stock, which will grow from that state. If there is an event of less fodder and water availability and farmers get reduced milk yields without loss of animal stock, then this will result in temporary effects on future production. However, due to less availability of fodder and water, farmers may sell some of their animals for slaughtering to make up their income losses. In this case, this will have permanent effect due to reduction in the state of animal stock.

For forecasting milk production, it must be identified whether the underlying process is a random walk model with drift or a trend-stationary process. Furthermore, this information would indicate whether the shocks have permanent or temporary effect on future production and thus may also be useful for policy makers in choosing the appropriate support services for the livestock sector.

There have been some efforts in previous studies on forecasting production of other agricultural products in Pakistan. A study on forecasting wheat production (Masood and Javed, 2004) used an autoregressive model in the analysis. An other study on forecasting citrus production and export estimated an Autoregressive Integrated Moving Average (ARIMA) model (Ahmad *et al*., 2005). These studies, however, did not compare the time series models to identify whether the production shocks have permanent or temporary effect on future production.

The objective of this study was to forecast milk production in Pakistan. First, we apply the Dickey-Fuller test for unit root to determine whether the time series of milk production is represented as a random walk model with drift or a trend-stationary process. Then, we use the accepted model for forecasting milk production for the next ten years.

Materials and Methods

Model and Estimation Method

Selection of a time series model for forecasting depends on the time series properties of the variable. When there is a trend in the data, the true process may be a random walk model with drift:

$$y_t = \alpha + y_{t-1} + \varepsilon_t \tag{1}$$

or it may be a trend-stationary autoregressive model:

$$y_t = \alpha + \delta t + \rho y_{t-1} + \varepsilon_t \text{ with } |\rho| < 1, \tag{2}$$

where y_t is a random variable, α, δ, ρ are parameters and ε_t is the error term, which is a white noise process.

The random walk model with drift in Eq. 1 is an autoregressive model containing unit root, as its coefficient on y_t is unity. The trend-stationary autoregressive model in Eq. 2 does not contain a unit root, as $|\rho|<1$. We apply the Dickey-Fuller test for unit root to determine which model is represented for the time series.

To carry out the Dickey-Fuller test, the general model for estimation is:

$$y_t = \alpha + \delta t + \rho y_{t-1} + \varepsilon_t, \tag{3}$$

which can be a random walk model with drift or a trend-stationary autoregressive model depending on the values of δ and ρ. If $\delta = 0$ and $\rho = 1$, then Eq. 3 is a random walk model with drift, otherwise, it is a trend-stationary autoregressive model. In the Dickey-Fuller test, the null hypothesis (H_0) is random walk model with drift and the alternative hypothesis (H_A) is trend-stationary autoregressive model. The F test is performed for testing the joint null hypothesis that $\delta = 0$ and $\rho = 1$ in Eq. 3. For performing the F test, Eq. 3 is estimated as unrestricted model under H_A and as restricted model with the restrictions of $\delta = 0$ and $\rho = 0$ under H_0. The ordinary least square (OLS) method is used for estimating the equation. The OLS F statistic is computed as follows:

$$F = \frac{(e_*'e_* - e'e)/(P - P_*)}{e'e/(N - P)}, \tag{4}$$

where e_* is vector of residuals from H_0, $e_*'e_*$ is residual sum of squares (RSS) from H_0 and P_* is its number of parameters. Similarly, e is vector of residuals from H_A, $e'e$ is RSS from H_A, P is its number of parameters and N is the number of sample observations. For the Dickey-Fuller test for unit root, the OLS F statistic is compared with the critical values provided by Dickey and Fuller (1981) as reported in Hamilton (1994). If the F statistic is larger than the critical value, the null hypothesis is rejected.

In Eq. 1 and 2, it is assumed that the error term ε_t is a white noise process. The Ljung and Box (1979) test is conducted for checking that the error term is a white noise process.

Data and Forecasting Method

Data used in this study are annual fresh milk production (million tonnes) in Pakistan from 1971-72 to 2004-05 published in Pakistan Economic Survey (Anonymous, 1998; 2005a), presented in Appendix A. For estimating Eq. 3, the data are transformed into natural logarithm:

$$y_t = \ln(M_t), \tag{5}$$

where M_t is quantity of milk produced in a year. This transformation is appropriate for milk production and many trended economic time series as the transformed model characterizes the growth in terms of rate, where as the linear model characterizes the growth in absolute terms. Given the conditional expected value $E[y_{t+1} \mid y_t]$ and the variance σ^2 and if we assume that the error term is normally distributed, the expected value and variance of M_{t+1} are computed as:

$$E[M_{t+1} \mid M_t] = e^{E[y_{t+1}|y_t] + \sigma^2/2} \tag{6}$$

$$\sigma_M^2 = e^{2E[y_{t+1}|y_t] + \sigma^2}\left(e^{\sigma^2} - 1\right) \tag{7}$$

The forecast of milk production M_{t+1} is computed using Eq. 6. The 95% confidence interval of forecasting is estimated as:

$$E\left[M_{t+1} \mid M_t\right] \pm 1.96\sqrt{MSE} \tag{8}$$

where MSE is mean squared error. The value of MSE depends on the estimated model. In this study, the random walk model with drift in Eq. 1 is accepted. In this case, its mean squared error is computed as:

$$MSE = s\sigma_M^2 \tag{9}$$

where s is number of years farther from the current year and σ_M^2 is given in Eq. 7.

Results and Discussion

Milk production in Pakistan has grown from 7.800 million tonnes in 1971-72 to 29.472 million tonnes in 2004-05. The positive trend in milk production may arise from the positive drift term of a random walk model or from a deterministic time trend of a trend-stationary autoregressive model. Table 1 presents the OLS estimates of both models. Ljung-Box test in both models indicated that the error term was a white noise process.

Random walk model with drift is a restricted model with the restrictions of $\delta = 0$ and $\rho = 1$. This model has a unit root as $\rho = 1$. The trend-stationary autoregressive model is unrestricted model. The Dickey-Fuller test is carried out to determine which model is represented for milk production. In this test, the null hypothesis is random walk model with drift. The alternative hypothesis is trend-stationary autoregressive model. For testing the null hypothesis, F statistic was computed using Eq. 4 and was equal to 3.42961. Results of the Dickey-Fuller test show that the null hypothesis is not rejected as the F statistic is much less than 7.24, which is the critical value at 5% significance level provided in Dickey and Fuller (1981). These results indicate that milk production is characterized as random walk model with drift. Thus, the shocks to milk production in a year have permanent consequences for the level of its future production.

Milk production is forecasted using the results of random walk model with drift. Table 2 presents annual milk production forecasts for the next ten years from 2005-06 to 2014-15. The expected milk production in the year 2005-06 is 30.7 million tonnes with the 95% confidence interval from 28.722 to 32.678 million tonnes. Similarly, the expected milk production and its confidence interval are presented for each year in the Table 2. The growth rate in the expected milk production is 4.17%/annum in each year from 2005-06 to 2014-15.

Table 1: Estimation of random walk model with drift and trend-stationary autoregressive model

		Random walk model with drift (H_0)			Trend-stationary autoregressive model (H_A)		
Regressor	Parameter	Estimate	t-statistic	p-value	Estimate	t-statistic	p-value
Constant	α	0.0403	7.0413	0.0000	0.3270	2.4914	0.0185
t	δ	0	--	--	0.0082	2.4677	0.0195
ln (M_{t-1})	ρ	1	--	--	0.8383	11.8421	0.0000
R^2		0.9949			0.9959		

Table 2: Forecasts of annual milk production in Pakistan

		95% Confidence interval	
Year	Milk forecast E [M] (million tonnes)	Lower (million tonnes)	Upper (million tonnes)
2005-06	30.700	28.722	32.678
2006-07	31.979	29.065	34.894
2007-08	33.312	29.592	37.031
2008-09	34.700	30.225	39.175
2009-10	36.145	30.932	41.359
2010-11	37.652	31.701	43.602
2011-12	39.220	32.524	45.917
2012-13	40.855	33.395	48.314
2013-14	42.557	34.313	50.801
2014-15	44.330	35.276	53.384

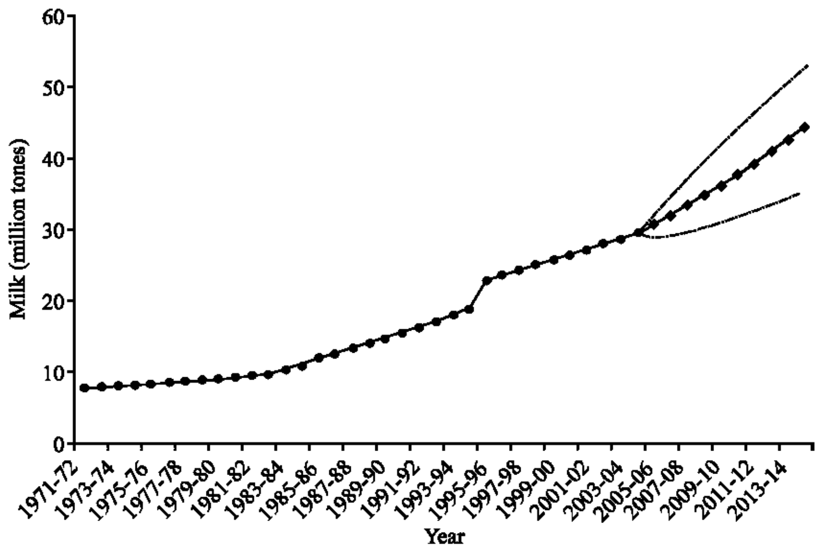

Fig. 1: Annual milk production and forecasts with confidence interval

Figure 1 presents the actual milk production from 1971-72 to 2004-05 (Appendix A) and the forecasted milk production from 2005-06 to 2014-15 with its 95 confidence interval (Table 2). As the accepted model is a random walk model with drift, the confidence interval expands as the forecast is done farther from the current year. This is because the mean squared error grows linearly with the forecast horizon.

Appendix A: Annual milk production in Pakistan from 1971-72 to 2004-05

Year	Milk production (million tones)
1971-72	7.800
1972-73	7.899
1973-74	8.044
1974-75	8.173
1975-76	8.348
1976-77	8.524
1977-78	8.704
1978-79	8.888

Appendix A: Continued

Year	Milk production (million tones)
1979-80	9.075
1980-81	9.267
1981-82	9.462
1982-83	9.662
1984-84	10.242
1984-85	10.856
1985-86	12.052
1986-87	12.669
1987-88	13.319
1988-89	14.003
1989-90	14.723
1990-91	15.481
1991-92	16.280
1992-93	17.120
1993-94	18.006
1994-95	18.986
1995-96	22.970
1996-97	23.580
1997-98	24.215
1998-99	24.876
1999-00	25.566
2000-01	26.284
2001-02	27.031
2002-03	27.811
2003-04	28.624
2004-05	29.472

Source: Pakistan economic survey (Anonymous, 2005a and 1998)

Conclusion

Milk production in Pakistan is forecasted using time series data from 1971-72 to 2004-05. As there is a trend in the data, the representative model may be the random walk model with drift or trend-stationary autoregressive model. The results of Dickey-Fuller test for unit root show that the time series of milk production is represented as a random walk model with drift, which indicates that the shocks to production in a year have permanent effect on the level of future production. The forecast results show the expected growth rate in milk production is 4.17%/annum. Although per capita availability is expected to grow as population growth rate in Pakistan is about 2%, there will be more demand for milk due to urbanization and higher income of households with economic growth.

References

Ahmad, B., A. Ghafoor and H. Baddar, 2005. Forecasting and growth trends of production and export of kinnow from Pakistan. J. Agric. Soc. Sci., 1: 20-24.

Anonymous, 1998. Economic Survey 1997-98. Government of Pakistan, Finance Division, Economic Advisor's Wing, Islamabad, June 1998.

Anonymous, 2001. Livestock Development: Implications for Rural Poverty, the Environment and Global Food Security. The World Bank, Washington, DC, USA, pp: 1-8.

Anonymous, 2004. National Accounts of Pakistan: Rebasing from 1980-81 to 1999-2000. Government of Pakistan, Statistics Division, Federal Bureau of Statistics, Islamabad.

Anonymous, 2005a. Pakistan Economic Survey 2004-05. Government of Pakistan, Finance Division, Economic Advisor's Wing, Islamabad, June 2005.

Anonymous, 2005b. FAO Statistics Database on the Worldwide Web: http://apps.fao.org (accessed in September 2005).

Dickey, D.A. and W.A. Fuller, 1981. Likelihood ratio statistics for autoregressive time series with a unit root. Econometrica, 49: 1057-72.

Hamilton, J.D., 1994. Time Series Analysis. Princeton University Press, Princeton, NJ, USA.

Ljung, G. and G. Box, 1979. On a measure of lack of fit in time series models. Biometrika, 66: 265-270.

Masood, M.A. and A. Javed, 2004. Wheat production forecast model for Pakistan. Pak. J. Agril. Econom., 5: 69-82.