Quest Journals
Journal of Research in Applied Mathematics
Volume 2~ Issue 9 (2016) pp: 07-09
ISSN(Online) : 2394-0743 ISSN (Print):2394-0735
www.questjournals.org

Research Paper

Some Aspects of Compound Geometric Lag Models in Agricultural Field

A.V.S.Raghavaiah[1] & Prof. P.Mohammed Akhtar[2]
[1]*Lecturer In Statistics, SKSC Degree College (Affiliated To Y V University), Proddatur, Kadapa (Dist)*
[2]*Professor, Department Of Statistics, S.K University, Anantapuramu.*

Received 01June, 2016; **A**ccepted 13 June, 2016 © The author(s) 2016. Published with open access at**www.questjournals.org**

ABSTRACT:-*Mathematics, Economics and Statistics are integrated for the purpose of providing estimation for the parameters of the economic relationships. Various Econometric methods are in vogue to estimate random disturbances or errors using different types of estimation procedures. In the present study from the existing inferential aspects of linear regression some new models are developed by using compound geometric lag (CGL) model. Regression coefficients bi (i= 1, 2, 3…k), R^2 and F are estimated from these models. Based on the results the significance of regression coefficients and R^2 are tested. Finally the best CGL model is to be determined.*

Key-words:- *Linear Regression model, CGL, regression co-coefficients, lag, R^2, F.*

I. INTRODUCTION

In India agriculture is given priority in every five year plan. The agriculture production especially in the food crops like paddy and ground nut is increased through intensive cultivation adopting modern agriculture technology. CGL model has become imperative for proper planning of agricultural activities for further increase of the production. Raghavaiah.A.V.S and Mohammed Akhtar.P [11] have developed a CGL model using data mining for the estimation of regression coefficients of the model and tested its goodness of fit for the proposed model of groundnut data.

In this paper an attempt is made to suggest a suitable model for the present study and examine the yield response analysis of groundnut in Anantapur Dist. of A.P

II. DATA MINING

Statistics has been used to analyze the data in an effort to find correlation, pattern and dependencies. However with an increased technology more and more data is available, for analyzing it manually is difficult. So computers are used for the analysis purpose. Data mining is a term used to describe the "Processing of discovery patterns and trends in large data sets in order to find useful decision making information". Data mining is commonly defined as the computer assisted search for interesting patterns, and relations in a large data bases. It is a relativity young area of research that built on the olden discipline of the statistics, data base, artificial intelligence and date visualization. Data mining is usually considered to be a form of secondary data analysis. This means that it is often performed on data collected for administrative purpose.

In this regard we consider possibilities are applied data mining in fitting of econometric models. We consider data mining as a collection of techniques and algorithms that have been developed in this area of research. Data mining involves scientists for wide range of disciplines, including Mathematics, Statistics, and computer professional etc..,

In this paper we consider a small part of Data mining from agriculture data to fit and test our proposed econometric models.

III. METHODOLOGY

A realistic formulation of economic relations often requires the introduction of lagged values of dependent variables as explanatory variables. These lags are of greater importance for decision making in the economic data. Farmers faced many problems on making decision on yield in response to various economic

factors. At present, the following Compound Geometric Lag Models are used to study the effect of risk factors on yield response.

$\mathbf{Y_t = b_0 + b_1 A_{t-1} + b_2 Y_{t-1} + b_3 Y_{t-2} + b_4 R_t + b_5 R_{t-1}}$…….(3.1)

Excluded R_{t-1} from model-1. We get,

$\mathbf{Y_t = b_0 + b_1 A_{t-1} + b_2 Y_{t-1} + b_3 Y_{t-2} + b_4 R_t}$ …….(3.2) Excluded R_t from model-1. We get,

$\mathbf{Y_t = b_0 + b_1 A_{t-1} + b_2 Y_{t-1} + b_3 Y_{t-2} + b_5 R_{t-1}}$……... (3.3)

Excluded Y_{t-2} from model-1. We get,

$\mathbf{Y_t = b_0 + b_1 A_{t-1} + b_2 Y_{t-1} + b_4 R_t + b_5 R_{t-1}}$…….. (3.4)

Excluded Y_{t-1} from model-1. We get,

$\mathbf{Y_t = b_0 + b_1 A_{t-1} + b_3 Y_{t-2} + b_4 R_t + b_5 R_{t-1}}$…….. (3.5)

Here A_t = Current hectare age under the crop**;**

A_{t-1} = Hectarage under the crop lagged by one year

Y_t = Yield per hectarage**;**

Y_{t-1} = Yield per hectarage lagged by one year

Y_{t-2} = Yield per hectarage lagged by two years

R_t = Rainfall in current year**;**

R_{t-1} = Rainfall lagged by one year.

A secondary data is taken from the Hand book of Statistics, Chief Planning Officer, Anantapur **Dist. A.P** of Groundnut crop with area, yield, and rainfall during year 1989-90 to 2011-2012 of Anantapur Dist. The following table –I gives the data on groundnut crop with various factors.

TABLE-I

Table 1: Area, yield of groundnut crop & rain fall during the year 1989-90 to 2011-12 of ANANTAPURAM

S.N.	Years	Y_t	A_t	A_{t-1}	Y_{t-1}	Y_{t-2}	R_t	R_{t-1}
1	1989-90	680	730	715	967	1023	705	757
2	1990-91	664	741	730	680	967	482	705
3	1991-92	689	721	741	664	680	574	482
4	1992-93	680	743	721	689	664	500	574
5	1993-94	1011	731	743	680	689	593	500
6	1994-95	621	687	731	1011	680	377	593
7	1995-96	1328	744	687	621	1011	531	377
8	1996-97	731	759	744	1328	621	750	531
9	1997-98	373	671	759	731	1328	441	750
10	1998-99	1156	781	671	373	731	695	441
11	1999-2000	383	717	781	1156	373	521	695
12	2000-2001	1116	815	717	383	1156	612	521
13	2001-2002	467	778	815	1116	383	702	612
14	2002-2003	372	750	778	467	1116	290	702
15	2003-2004	303	686	750	372	467	523	290
16	2004-2005	810	872	686	303	372	434	523
17	2005-2006	436	899	872	810	303	791	434
18	2006-2007	94	662	899	436	810	408	791
19	2007-2008	1260	897	662	94	436	816	408
20	2008-2009	115	871	897	1260	94	714	816
21	2009-2010	268	530	871	115	1260	616	714
22	2010-2011	557	834	530	268	115	722	616
23	2011-2012	275	754	834	557	268	493	722

$Y_t = 2163.344 - 2.003A_{t-1} + 0.089y_{t-1} + 0.358y_{t-2} + 0.544R_t - 1.090R_{t-1}$

The data given in table-I is taken and fitted to the models (3.1) to (3.5). We get the following results given in table-II.

TABLE-II

Model	R^2	F	constant	A_{t-1}	Y_{t-1}	Y_{t-2}	R_t	R_{t-1}
			b_0	b_1	b_2	b_3	b_4	b_5
1	0.699	7.913*	2163.34	-2.003 *	0.089 *	0.358 *	0.544 *	-1.090 *
2	0.556	5.64 *	1928.63	-2.5980 *	-0.016 *	0.3077 *	0.788 *	-
3	0.6622	8.82 *	2616.64	-2.05 *	0.118 *	0.291 *	-	-1.22 *
4	0.5922	6.535 *	2554.76	-1.99 *	0.011 *	-	0.2516 *	-0.976 *
5	0.6932	10.17 **	2119.87	-1.92 *	-	0.339 *	0.574 *	-1.034 *

Significant at 1% probability of F (9.68)
** Not Significant at 1% probability of F (9.68)

The results explained in table-1I relate to Anantapur Dist. of A.P for the crop of Groundnut.

- The first model contains all the variables, i.e., the dependent variables is yield per hectarage under the crop in the current year (Y_t) while independent variables are lagged hectarage (At_{-1}), lagged yields Y_{t-1}, $Y_{t-2,}$ current rain fall (R_t) and lagged rainfall R_{t-1} In this model the values of the risk factors or regression co-efficient $b_{1,}$ b_5 are negative while $b_{2,}$ $b_{3,}$ $b_{4,}$ are positive. . The values of R^2 is calculated as 69.9%. F-Test is used and its value is found to be 7.913 < 9.68 at 1%. So it is significant.
- The second model contains all variables except lagged rainfall (R_{t-1}).In this model the value of risk factor b_1, b_2 are negative, while $b_{3,}$ $b_{4,}$ are positive. The value of R^2 is calculated as 55.6%. F- Test is used and its value is found to be 5.64 < 9.68 at 1%. So the value is considered significant.
- The third model contains all variables of the first model except current rainfall (R_t).The values of the risk factors $b_{1,}$ b_5 are negative while $b_{2,}$ $b_{3,}$ are positive. The values of R^2 are calculated as 66.22%. F -Test was used and find its value as 8.82 < 9.68 at 1%. So the value is significant.
- In the fourth model contain all variables of first model except lagged yield by two years (Y_{t-2}). In this model the risk factor b_1, b_5 are negative while $b_{2,}$ b_4 are positive. The results of R^2 are calculated as 59.22%. F-Test is used and its value is 6.535 < 9.68 at 1%. So the value is significant.
- The fifth model contains all variables of the first model except lagged yield by one year (Y_{t-1}). In this model the risk factor b_1, b_5 are negative while $b_{3,}$ b_4 are positive. The results of R^2 are explained about 69.32%. F-Test is used and its value is found to be 10.17> 9.68 at 1%. So the value is not significant.
- Generally the higher R^2 value is selected as the best model.
- I with higher R^2 value (69.9%) is selected to be fit for the data for drawing inference.

IV. Conclusion

- The co-efficient of lagged hectare and lagged rainfall are negative in all models and are significant.
- The coefficients of lagged yield by two year are positive in all models, and are significant.
- The co efficient of lagged yields by one year is positive in all models except in model-II are significant.
- The price factor is also playing a vital role to allocate the land under a particular crop. So, we introduced current price (P_t) and lagged prices (P_{t-1}) in the proposed model and find the effect in risk factor of prices.
- Anantapur district has been sealing under frequent droughts and famines with scanty rainfall. Agriculture being a gamble with rain in the district, people have to be extremely cautious in allocating land for different agricultural operations. The study helps the farmers immensively to allocate land for cultivating groundnut.

REFERENCES

[1]. Gorman,J.W. and Toman,R.J.(1966),"Selection of Variables For fitting Equations to data"Technometrics,8,27-51.
[2]. Huang,D.Y.(1966),"Selection procedures in linear models", Journal of Statistical planning and inference,54,271-277
[3]. Pesaran, M.H.(1974),"On the problem of Model selection", Review of Economic Studies,41,153-171.
[4]. Hocking, R.R (1976)."The Analysis and selection of variables in Linear Regression", Biometrics, 32, 1-4.
[5]. Lovell, Mc,C (1983),"DataMining",Review of Economics and Statistics" ,Vol.65, pp-112.
[6]. Ullah,A.(1985), "Specification Analysis of Econometric Models", Journal of
[7]. Quantitative Economics,1,pp-187-209.
[8]. Jonston,J., and Dinardo,J.(1997), "Econometric Methods" Fourth Edition,Mc Grew Hill Company, NewYark.George, I.E.(2000), "The variable selection problem",V-95, pp-1304-1308.
[9]. Sargan, J.D. (2001), "Model building and mining", Economic review, V-20. 159-170. Anderson.T.W. Apendix10.1 "An Introduction to Multivariate Statistical Analysis"
[10]. Raghavaiah.A.V.S. & Mohammed Akhtar.P. (2016) " Estimation of Regression
[11]. Co efficient through Compound Geometric Lag Model using Data Mining" Bulletin of Pure and Applied Sciences, Volume 35E (Maths & Stat.) Issue (No.1) 2016: P.1-11.