P: ISSN NO.: 2394-0344 RNI No.UPBIL/2016/67980 VOL-3* ISSUE-4*(Supplementary Issue) July- 2018
E: ISSN NO.: 2455-0817 Remarking An Analisation

Nonstationary Scalar Time Series

Abstract

In this paper, a time series {X(t, ω), t Є T } where X is a random variable (r. v.) on (Ω, C, P) is explained. The properties of non-stability with supporting real life examples have been taken and conclusions have been drawn by testing methodology of hypothesis. Area of Kharif jawar data for 27 years from five districts of Marathwada of Maharastra State were analyzed.

A preliminary discussion of properties of time series precedes the actual application to regional district-wise area of kharif jawar data.

Keywords: Time Series, Regression Equation, Auto-Covariance, Auto-Correlation.

B. L. Bable
Head,
Department of Statistics,
Dnyanopasak Mahavidyalaya,
Parbhani, Maharastra

D. D. Pawar
Professor and Chairman,
Department of Statistics,
N.E.S. Science College,
Nanded, Maharastra

Introduction

Our aim here is to illustrate a few properties of time series with supporting real life examples. Concepts of auto covariance and auto correlation are shown to be useful which can be easily introduced. In this article we have used area of kharif jawar data of 1976 to 2002 at five locations in Marathwada region to illustrate most of properties theoretically established.

Objectives of the Study

1. To develop theory of time series. Specially improving the theorems which characterize time series.
2. To develop algorithms for analyzing time series, which use the characterizing theorems.
3. By using data from Marathwada Region for validating the algorithms, and testing the methods.
4. To interpret the results of characterizations, in real economic and social terms.

The main purpose of this work is to summarize the research work carried out on the above given objectives and to draw useful conclusions on the basis of auto regressive time series analysis. A way to check trends and randomness in the data scalar time series by using properties of auto covariance.

Basic Concepts

Basic definitions and few properties of stationary time series are given in this section.

Definition 2.1: A Time Series

Let (Ω, C, P) be a probability space let T be an index set. A real valued time series is a real valued function X (t ,ω) defined on T x Ω such that for each fixed t Є T, X(t, ω) is a random variable on (Ω, C, P).

The function X(t, ω) is written as X(ω) or X_t and a time series considered as a collection {X_t : t Є T }, of random variables [12].

Definition 2.3: Stationary Time Series

A process whose probability structure does not change with time is called stationary. Broadly speaking a time series is said to be stationary, if there is no systematic change in mean i.e. no trend and there is no systematic change in variance.

Definition 2.4: Strictly stationary time series

A time series is called strictly stationary, if their joint distribution function satisfy,

$$F_{X_{t1} X_{t2} \dots X_{tn}}(x_{t1}, x_{t2} \dots x_{tn}) = F_{X_{t1+h} X_{t2+h} \dots X_{tn+h}}(x_{t1}, x_{2t} \dots x_{tn}) \quad \dots (1)$$

Where, the equality must hold for all possible sets of indices ti and (ti + h) in the index set. Further the joint distribution depends only on the distance h between the elements in the index set and not on their actual values.

Theorem 2.1

If { X_t : t Є T }, is strictly stationary with E{ $|X_t|$ } < α and E{ $|X_t - \mu|$ } < β then,

$E(X_t) = E(X_{t+h})$, for all t, h and } ….(2)
$E[(X_{t_1} - \mu)(X_{t_2} - \mu)] = E[(X_{t_{1+h}} - \mu)(X_{t_{2+h}} - \mu)]$, for all t_1, t_2, h

Proof

Proof follows from definition (2.4).

In usual cases above equation (2) is used to determine that a time series is stationary i.e. there is no trend.

Definition 2.5: Weakly Stationary Time Series

A time series is called weakly stationary if
1.The expected value of X_t is a constant for all t.
2.The covariance matrix of $(X_{t1}, X_{t2}, \ldots X_{tn})$ is same as covariance matrix of $(X_{t1+h}, X_{t2+h}, \ldots X_{tn+h})$.

A look in the covariance matrix $(X_{t1} X_{t2} \ldots X_{tn})$ would show that diagonal terms would contain terms covariance (X_{ti}, X_{ti}) which are essentially variances and off diagonal terms would contains terms like covariance (X_{ti}, X_{tj}). Hence, the definitions to follow assume importance. Since these involve elements from the same set $\{X_{ti}\}$, the variances and co- variances are called auto-variances and auto-co variances .

Definition 2.6: Auto-covariance function: The covariance between $\{X_t\}$ and $\{X_{t+h}\}$ separated by h time unit is called auto-covariance at lag h and is denoted by $\Upsilon(h)$.

$\Upsilon(h) = cov(X_t, X_{t+h}) = E\{X_t - \mu\}\{X_{t+h} - \mu\}$ ...(3)

The function $\Upsilon(h)$ is called the auto covariance function.

Definition 2.7: The Auto Correlation Function

The correlation between observation which are separated by h time unit is called auto-correlation at lag h. It is given by

$$P(h) = \frac{E\{X_t - \mu\}\{X_{t+h} - \mu\}}{[E\{X_t - \mu\}^2 E\{X_{t+h} - \mu\}^2]^{1/2}} \quad \ldots(4)$$

$$= \frac{\Upsilon(h)}{[E\{X_t - \mu\}^2 E\{X_{t+h} - \mu\}^2]^{1/2}}$$

Where μ is mean.

Remark 2.1

For a vector stationary time series the variance at time (it + h) is same as that at time it. Thus, the auto correlation at lag h is

$$\rho(h) = \frac{\Upsilon(h)}{\Upsilon(0)} \quad \ldots(5)$$

Remark 2.2

For h = 0, we get, $\rho(0) = 1$.

For application attempts have been made to establish that area at certain districts of Marathwada satisfy equation (1) and (5).

Theorem 2.2: The covariance of a real valued stationary time series is an even function of h.

i.e., $\Upsilon(h) = \Upsilon(-h)$.

Proof

We assume that without loss of generality, $E\{X_t\} = 0$, then since the series is stationary we get, $E\{X_t X_{t+h}\} = \Upsilon(h)$, for all t and t + h contained in the index set. Therefore if we set $t_0 = t_1 - h$,

$\Upsilon(h) = E\{X_{t0} X_{t0+h}\} = E\{X_{t1} X_{t1+h}\} = \Upsilon(-h)$ …..(6)

Proved.

Theorem 2.3

Let X_t's be independently and identically distributed with $E(X_t) = \mu$ and $var(X_t) = \sigma^2$
then $\Upsilon(t, k) = E(X_t, X_k) = \sigma^2$, $t = k$ $= 0$, $t \neq k$

This process is stationary in the strict sense.

Testing Procedure

Inference Concerning Slope (β_1)

For testing $H_0: \beta_1 = 0$ Vs $H_1: \beta_1 > 0$ for $\alpha = 0.05$ percent level using t distribution with degrees of freedom is equal to n – 2 were considered.

$$t_{n-2} = \beta_1 / s_{\beta 1}$$

where β_1 is the slope of the regression line and $s_{\beta1} = s_e / s_t$ and $s_e = [SSE / n - 2]^{1/2}$,

sum of squares due to errors (SSE) = $(s_t^2 - s_{tx}^2 / s_t^2)$, $s_{tx} = \Sigma(t_i - \bar{t})(X_t - \bar{X})$

$s_t^2 = \Sigma(t_i - \bar{t})^2$; $s_x^2 = \Sigma(X_t - \bar{X})^2$

.Where SSE is the sum of squares due to error or residual sum of squares .

Example of Time Series

Area of Kharif jawar data of Marathwada region were collected from five districts namely Aurangabad, Parbhani, Beed, Osmanabad, and Nanded. The data were collected from Socio Economic Review and District Statistical Abstract, Directorate Economics and Statistics Government of Maharastra Bombay, Maharastra Quarterly Bulletin of Economics and Statistics, Directorate of Economics and Statistics Government of Maharastra, Bombay [2, 3 , 4]. Hence.we have five dimensional time series t_i, i = 1, 2 , 3 , 4 , 5 corresponding to the districts Aurangabad, Parbhani, Beed, Osmanabad and Nanded respectively. Table 4.1A, shows the results of descriptive statistics, table 4.1B and table 4.2C shows linear trend analysis. All the linear trends were found to be not significant except Parbhani and Aurngabad districts.

Over the years many scientists have analyzed rainfall, temperature, humidity, agricultural area, production and productivity of region of Maharastra state, [1 , 5 , 6, 7 , 8, 9, 11 , 14, 15] . Most of them have treated the time series for each of the revenue districts as independent time series and tried to examine the stability or non-stability depending upon series. Most of the times non- stability has been concluded and hence possibly any sort of different treatment was possibility never thought of. In this investigation we treat the series first and individual series. The method of testing intercept ($\beta_0 = 0$) and regression coefficient ($\beta_1 = 0$), Hooda R. P. [13] and for testing correlation coefficient Bhattacharya G.K. and Johanson R.A. [10]. We set up null hypothesis for test statistic used to test, $H_0 : \beta_1 = 0$ and $H_1 : \beta 1 > 0$, for $\alpha = 0.05$,

-------- --------

$t_{n-2} = \beta \sqrt{S_{xx}} / \hat{\sigma}$, where $\hat{\sigma} = \sqrt{SSE / n-2}$

The hypothesis H_0 is not significant for both the values of t for 25and 19 d. f. for each districts.

The regression analysis tool provided in MS-Excel was used to compute β_0 , β_1, corresponding SE , t-values for the coefficients in regression models. Results are reported in table 4.1B and table4.2C. Elementary statistical analysis is reported in table–4.1A. It is evident from the values of CV that there is hardly any scatter of values around the mean indicating that all the series are not having trend.

Table4.1B shows that the model,

$$X_t = \beta_0 + \beta_1 t + \in,$$

When applied to the data indicates $H_0 : \beta_1 = 0$ is true . Hence X_t is a not having trend for four districts except Parbhani and Aurngabad districts.

$$X = \beta_0 + \beta_1 t + \in,$$

where,

1. X_t are the area of kharif jawar series.
2. t is the time (years) variable.
3. ε is a random error term normally distributed as mean 0 and variance σ^2 .

Area X_t in is the dependent variable and time t in (years) is the independent variable.

Values of auto covariance computed for various values of h are given in table-4.2A. values for different districts were input as a matrix to the software. Defining

$A = y_1 , y_2 \ldots\ldots y_{n-h}$

$B = y_{h+0} , y_{h+1} \ldots\ldots y_n$

Υ (h) = cov (A , B) were computed for various values of h. Since the time series constituted of 33 values, at least 10 values were included in the computation. The relation between Υ(h) were examined using model, table-4.2C.

$\Upsilon(h) = \beta_0 + \beta_1 h + \in$,

The testing shows that , both the hypothesis $\beta_0 = 0$ and $\beta_1 = 0$ test is positive . Table-4.2C was obtained by regressing values of Υ (h) and h , using “Data Analysis Tools” provided in **MS Excel**. Table 4.2A formed the input for table 4.2C. In other wards, Υ (h) are all zero except Parbhani and Aurngabad districts, in the area series of Parbhani and Aurngabad districts trend was found showing that X_t , X_{t+h} are dependent in area series of Parbhani and Aurngabad districts and there is a trend in that series. Hence in Parbhani and Aurngabad districts area X_t is not stationary it presents a trend.

Conclusion

It was observed that t values are therefore significant for the Parbhani district, rest of the four districts are not significant i.e. concluded that X_t does not depend on t for 4 districts [5] . Similarly, $\Upsilon_{i j}$(h) does not depend on h to mean that , ‘no linear relation’ rather than ‘no relation in four districts except Aurngabad district. The testing shows that, for the hypothesis $\beta_1 = 0$, test is positive for t for 4 district.

Generally it is expected, area of kharif jawar over a long period at any region to be not stationary time series. These results does not conform with the series in Aurngabad district i,e. in Aurngabad district trend was found in area of kharif jawar series.

Analysis

Area of Kharif Jawar

The same strategy of analyzing first individual time series as scalar series and then treating the vector series as the regional time series has been adapted here for area series.

Area Time Series Treated As Scalar Time Series

The model considered was:

$X_i(t) = (\beta_0)_i + (\beta_1)_i t + \in_i(t)$, i =1, 2,…5 …(7)

Where X_i is the annual area series, t is the time series variable , β_0 = the intercept, β_1 = the slope, ϵ_i is the random error. Area X_i is the dependent variable and time t in years is the independent variable.

Table-4.1

Districtwise Area of Hundred Hectores of Kharif Jawar Time Series Data of Five Stations (Districts) in Marathwada Region

Sr. No.	Districts→ Years ↓	Aurangabad	Parbhani	Osmanabad	Beed	Nanded
1	1976	766	1583	2170	824	2262
2	1977	714	1828	2481	1145	2494
3	1978	723	1722	2276	828	2616
4	1979	503	919	1683	432	2402
5	1980	739	1609	2605	794	2500
6	1981	735	1132	2848	1062	2405
7	1982	735	1732	2605	1062	2405
8	1983	1087	2012	2848	1345	2535
9	1984	1087	2012	2848	1345	2535
10	1985	1116	1894	2938	1489	2439
11	1986	1121	2001	3171	1354	2502
12	1987	1112	1841	3159	1363	2426
13	1988	1314	1736	3251	1445	2554
14	1989	1222	1711	3120	1296	2509
15	1990	1268	1792	3348	1175	2513
16	1991	1307	1800	3252	1121	2507

17	1992	1248	1865	3264	1145	2617
18	1993	1365	1918	3150	1221	2613
19	1994	1420	1959	3208	1265	2733
20	1995	1179	2094	2614	1078	2315
21	1996	1268	2115	2666	1182	2481
22	1997	1286	2194	3059	1117	2322
23	1998	1163	2173	3034	1006	2332
24	1999	1169	2242	3339	1171	2550
25	2000	1061	1944	2707	1086	2374
26	2001	931	1853	2493	815	2196
27	2002	857	1902	2515	853	2160

Table-4.1A: Elementary Statistics of Area of Kharif Jawar Time Series Data of Marathwada Region For 27 Years (1976-2002).

Cities:	**Aurangabad**	**Parbhani**	**Osmanabad**	**Beed**	**Nanded**
Mean:	1123.0	1901.6	2934.2	944.3	2484.8
S.D.:	425.2	808.9	1147.6	389.4	1052.1
C.V.:	37.8	42.5	39.1	41.2	42.3

Table-4.1B: Linear Regression Analysis of Area of Kharif Jawar Time Series Data To Determine Trend Eq(7)

District	Coefficients		Standard Error	t Stat	Significance
Aurangabad	β_0	1230.27	173.22	7.10	**S**
	β_1	-7.66	10.81	-0.71	NS
Parbhani	β_0	1236.49	295.91	4.18	**S**
	β_1	47.51	18.47	**2.57***	**S**
Osmanabad	β_0	2359.54	453.42	5.20	**S**
	β_1	41.05	28.30	1.45	NS
Beed	β_0	1045.59	158.51	6.60	**S**
	β_1	-7.24	9.89	-0.73	NS
Nanded	β_0	1964.20	416.08	4.72	**S**
	β_1	37.19	25.97	1.43	NS

*t =2.06 is the critical value for 25 d f at 5% L. S. * shows the significant value*

A look at the table 4.1A shows that all of them have similar values of CV. Which indicates that their dispersion is almost identical. Trends were found to be not significant in 4 districts but ***significant*** in Parbhani district only. A simple look at the mean values shows that a classification as

C1 = {Aurangabad, Beed}

C2 = {Nanded, Parbhani, Osmanabad}

could be quite feasible.

In absence of linear trend, with reasonably low CV values can be taken as evidence of series being stationary series individually in four districts.

Further search for evidences of stability included determination of auto covariance and their dependency on lag variable h (Table 4.2A). Such an analysis requires an assumption of AR(Auto-regressive) model [12] Eq(8). Therefore a real test for stationary property of the time series can come by way of establishing auto- covariance's which do not depend on the lag variable

$X_t = C + \Phi X_{t-h} + \epsilon_t$, h = 0, 1, 2, … .20 … (8)

Table-4.2A

Auto Variances: Individual Column Treated As Ordinary Time Series For Lag Values (H = 0 , 1 , 2 ,………20) About Area Of Kharif Jawar Time Series Data.

lag h	**Aurangabad**	**Parbhani**	**Osmanabad**	**Beed**	**Nanded**
0	69429.7	20218.8	82558.2	55216.0	50611.5
1	57250.7	13087.5	53675.2	46569.9	36563.6
2	50994.5	8696.6	37012.7	42318.7	35355.7
3	46022.8	5659.6	18645.5	39243.3	30803.5
4	34618.6	1416.2	13681.9	32981.3	26713.6
5	23080.6	-4795.5	25848.7	26482.0	21339.7
6	12018.6	-7324.0	23834.0	23244.3	19180.3
7	481.1	-10180.4	15369.6	19153.0	13019.3
8	-10111.6	-12871.0	-5271.8	15493.8	7129.3
9	-16788.8	-13574.0	-22321.2	16111.7	1664.4
10	-23111.3	-10993.5	-20746.3	15992.9	4083.0
11	-31769.9	-5766.2	-9617.4	14660.3	670.8
12	-33484.6	514.2	-7680.2	10367.0	-8439.5
13	-29965.6	4180.2	1800.9	10644.1	-1095.6
14	-37804.5	3842.4	-25152.6	6775.2	-12775.8

15	-33725.6	4059.0	-38876.7	5434.4	-5145.1
16	-29199.0	4390.8	-32983.4	1381.3	-4616.7
17	-28931.7	332.7	-28586.3	-5250.8	-3003.3
18	-22766.1	-217.1	51.7	-5961.3	-3030.9
19	-18331.4	3768.4	12296.8	-1402.2	7.6
20	-15370.2	2078.8	1340.6	-5332.6	-3361.0

Table-4.2B : Correlation coefficient between h and Auto covariance is :

Corr. Coefficient	**-0.600***	-0.393	-0.254	-0.272	-0.334

*Correlation coefficient r = 0.433 is the critical value for 19 d f at 5% L. S. * shows the significant value.*

Correlation's between Υ_{ij} **(h)** and **h** were found ***significant*** in Auranga***bad*** district only showing that the time series can be reasonably assumed to be **not stationary** i.e. having trend**.** The coefficient is significant, with negative value showing that Aurangabad has been experiencing significantly declining area over the past years.

Table-4.2C: Linear Regression Analysis of Lag Values Vs Covariance.

District	Coefficients		Standard Error	t Stat	Significance
Aurangabad	β_0	51752.18	22999.26	2.25	**S**
	β_1	-6428.57	1967.35	**-3.27***	**S**
Parbhani	β_0	199290.26	87437.44	2.28	**S**
	β_1	-13926.87	7479.39	-1.86	NS
Osmanabad	β_0	236429.84	179714.73	1.32	NS
	β_1	-17627.05	15372.78	-1.15	NS
Beed	β_0	25104.43	22063.73	1.14	NS
	β_1	-2329.56	1887.33	-1.23	NS
Nanded	β_0	227749.09	142295.87	1.60	NS
	β_1	-18802.22	12171.97	-1.54	NS

*t =2.93 is the critical value for 19 d f at 5% L. S. * shows the significant value*

Acknowledgement

Authors are grateful to Prof. H.S. Acharya, Puna College, PUNE, for his valuable guidance.

References

1. *Acharya, H. S., Quantification of Climatic Component of Food Grains Production; J. of Maharastra Agric. Univ., 16(3) (1991), 324-326.*
2. *Anonymous, "Socio Economic Review and District Statistical Abstract", Directorate of Economic and Statistics, Government of Maharastra, Bombay (1970-1997)*
3. *Anonymous, "Maharastra Quarterly Bulletin of Economics and Statistics", Directorate of Economics and Statistics Govt. of Maharastra, Bombay (1998).*
4. *Anonymous, "Hand Book of Basic Statistics", Maharastra State (1999-2002).*
5. *Bable, B. L. and Acharya H.S., Time Series: The Concept and Properties to the Stationary Time Series; Bulletin of the Marathwada Mathematical Society., 7(1) (2006), 1-10.*
6. *Bable, B. L. and Pawar D., D.. Covariance Stationary Time Series; Multidisciplinary Bi-lingual International Journal 'Innovation'., 2(7) (2017), 28-32.*
7. *Bable, B. L. and Pawar D., D.. Properties of Weak and Strict Stationary Time Series;. Multidisciplinary International Research Journal Asian Resonance, 4(3) (2015), 54-58.*
8. *Bable, B. L. and Pawar D., D.. Trends in Scalar Time Series;. Multidisciplinary International Research Journal Periodic Research, 4(1) (2015), 41-46.*
9. *[9] Bable, B. L. and Pawar D., D.. Formation of Trends in Scalar Time Series;. International Research Journal Agricultural Economics and Statistics, 5(1) (2014), 111-115.*
10. *Bhattacharya, G. K. and Johanson R. A., Statistical Concepts and Methods. John Wiley and Sons, New York (1977).*
11. *Das, P. K., Statistics and Its Applications to Meteorology. Indian J. pure appl. Math., 26(6) (1995) , 531-546.*
12. *Fuller W. A., Introduction to Statistical Time Series. John Wiley and Sons, New York (1976).*
13. *Hooda R. P., Statistics for Business and Economics. Machmillan India Ltd., New Delhi (2003).*
14. *Patil, V. D. and Kalyanker S. P., Impact of Research on Productivity of Agricultural in Maharastra. Agri. Situ. India., 47(5) (1982), 321-326 .*
15. *Waghmare, P. R. and Katapallewar B. N., Districtwise trend in area, production and productivity of oilseeds in Maharastra State, 1980-81 to 1989-90. Indian J. Agric. Econ. , 48(3) (1993), 425-426.*