

Chi-Square Test in Time Series Data

Kelechukwu C. N. Dozie^{a*} and Stephen O. Ihekuna^a

^a Department of Statistics Imo State University, Owerri, Imo State, Nigeria.

Authors' contributions

This work was carried out in collaboration between both authors. Both authors read and approved the final manuscript.

Article Information

DOI: 10.9734/AJPAS/2022/v20i3426

Open Peer Review History:

This journal follows the Advanced Open Peer Review policy. Identity of the Reviewers, Editor(s) and additional Reviewers, peer review comments, different versions of the manuscript, comments of the editors, etc are available here: <https://www.sdiarticle5.com/review-history/90872>

Received 21 June 2022

Accepted 25 August 2022

Published 19 October 2022

Original Research Article

Abstract

This study examines the application of Chi-Square test in time series data which considers the mixed model structure and linear trending curve. The Chi-Square test is applied to the seasonal variances of the Buys-Ballot table. The emphasis is to assess the validity of the Chi-Square test using empirical examples. One hundred simulated numerical examples are used to illustrate the applicability of the Chi-Square test. Using empirical example, the Chi-Square test successfully recorded 100% of times for the mixed model. This expresses a high degree of confidence in the Chi-Square test.

Keywords: Buys-Ballot method; descriptive time series; mixed model; linear trend, seasonal variance; choice of model.

1 Introduction

The purpose of descriptive time series analysis is to isolate the four time series components available in the series. That is to de-compose an observed series ($X_t, t = 1, 2, \dots, n$) into components consisting the trend (T_t), the seasonal (S_t), the cyclical (C_t) and irregular (e_t) (Kendal and Ord [1], Chatfield [2]).

*Corresponding author: Email: kcndozie@yahoo.com;

The decomposition models are the additive, multiplicative and mixed models. For short series, the cyclical component is superimposed into trend (Chatfield [2]) and observed series $(X_t, t = 1, 2, \dots, n)$ can be decomposed into the trend-cycle component (M_t) , seasonal component (S_t) and irregular component (e_t) . Hence, the decomposition models are

$$\text{Additive Model } X_t = M_t + S_t + e_t \tag{1}$$

$$\text{Multiplicative Model } X_t = M_t \times S_t \times e_t \tag{2}$$

$$\text{Mixed Model } X_t = M_t \times S_t + e_t \tag{3}$$

It is assumed that the seasonal indices, when exist, has period s , that is it repeats after s time periods.

$$S_{t+s} = S_t, \text{ for all } t \tag{4}$$

For equation (1), the assumption is that, the sum of seasonal indices over a complete period is zero, ie.

$$\sum_{j=1}^s S_{t+j} = 0 \tag{5}$$

Again, for equations (2) and (3), the assumption is that, the sum of seasonal indices over a complete period is s .

$$\sum_{j=1}^s S_{t+j} = s \tag{6}$$

The assumption is that, the irregular component e_t is the Gaussian $N(0, \sigma_1^2)$ white noise for equations (1) and (3), while for equation (2), e_t is the Gaussian $N(1, \sigma_2^2)$ white noise and that $Cov(e_t, e_{t+k}) = 0, \forall k \neq 0$

One greatest problems identified in the use descriptive time series analysis is choice of suitable model for decomposition of any study series. That is when to use any of the three time series model is uncertain. It is important to note that, use of wrong model will definitely lead to erroneous estimate of the components.

To select among additive, multiplicative and mixed models, many scholars have suggested different approaches. Puerto and Rivera [3] proposed the use of coefficient of variation of seasonal differences $CV(d)$ and seasonal quotient $CV(c)$ for choice of model. According to them, additive model is appropriate, if $CV(c)$ is greater than $CV(d)$ and it is multiplicative if $CV(c)$ is less than or equal to $CV(d)$. However, they did not provide any statistical test to justify the use. Chatfield [2] suggested the use of time plot to choose between additive and multiplicative models. However, no theoretical basis was proposed for the decision rule. Iwueze, et al. [4] proposed the use of the relationship between the seasonal average $(\bar{X}_{.j}, j = 1, 2, \dots, s)$ and the seasonal standard deviations $(\hat{\sigma}_{.j}, j = 1, 2, \dots, s)$ to choose the appropriate model for decomposition.

In proposing the Chi-Square test Nwogu, et al. [5] and Dozie, et al [6] assumed that (i) the underlying distribution of the variable, $X_{ij}, i = 1, 2, \dots, m, j = 1, 2, \dots, s$, under study is normal. (ii) the observations in each column, $X_{ij}, i = 1, 2, \dots, m$ are independent and (iii) that the s -column are independent. Therefore, they

proposed use of Chi-Square test as a basis for choosing between mixed and multiplicative models. The seasonal variance for the mixed model is $\sigma_{.j}^2 = \frac{b^2 n(n+s)}{12} S_j^2 + \sigma_1^2$. Therefore, the test for choice between mixed and multiplicative model may be reduced to identify the mixed model whose seasonal variance is simply constant multiply of the square of seasonal effect.

The aim of this article is to assess the validity of the Chi-Square test using empirical examples. The rationale of the article is that it will help to improve existing methodology by providing analyst with objective Chi-Square test in a series when it exists. The article is limited to series when trend cycle component is linear.

2 Methodology

Chi-Square test is applied to the seasonal variances of the Buys-Ballot table and the model structure is mixed. Hence, the summary of the seasonal variance of the Buys-Ballot table for the mixed model derived by Nwogu, et al. [5] and Dozie, et al. [6] is shown in Table 1.

Table 1.

Linear trending curve	Seasonal Variance ($\sigma_{.j}^2$)
$a + bt$	$\frac{b^2 n(n+s)}{12} S_j^2 + \sigma_1^2$

In this arrangement, n is the total number of observations, s is the seasonal lag (number of columns), b is the slope, s_j is the seasonal indices, σ_1^2 is the error variance, assumed equal to 1. Both the b and s_j are derivable from the seasonal average.

$$\bar{X}_{.j} = \left[a + b \left(\frac{n-s}{2} \right) + b_j \right] * S_j \tag{7}$$

$$\equiv [\alpha + \beta_j] \times S_j \tag{8}$$

where,

$$\alpha = a + b \left(\frac{n-s}{2} \right), \beta = b \tag{9}$$

Estimates of α and β are derivable from the regression of $\bar{X}_{.j}$ on j and estimates of S_j is

$$\hat{S}_j = \frac{\bar{X}_{.j}}{\hat{\alpha} + \hat{\beta}_j} \tag{10}$$

For details of Buys-Ballot procedure, see Iwueze et.al [4], Nwogu et al. [5], Dozie et al. [6], Akpanta and Iwueze [7], Dozie and Ijeomah [8], Dozie and Nwanya [9], Dozie [10], Dozie and Uwaezuoke[11], Dozie and Ibebuogu [12], Dozie and Ihekuna [13] .

2.1 Chi-Square Test

The expression of the seasonal variance for the mixed model is $\sigma_{zj}^2 = \frac{b^2 n(n+s)}{12} S_j^2 + \sigma_1^2$. Therefore, Chi-Square test proposed by Nwogu, et al. [5] and Dozie, et al. [6] when trend-cycle component is linear reduces to that of test null hypothesis.

$$H_0 : \sigma_j^2 = \sigma_{zj}^2$$

and the suitable model is mixed

$$H_1 : \sigma_j^2 \neq \sigma_{zj}^2$$

and the suitable model is not mixed

$\sigma_j^2 = (j = 1, 2, \dots, s)$ is the true variance of the j th season.

$$\sigma_{zj}^2 = \frac{b^2 n(n+s)}{12} S_j^2 + \sigma_1^2 \tag{11}$$

and σ_1^2 is the error variance assumed to be equal to 1

$$\text{Therefore, the statistic is } \chi_c^2 = \frac{(m-1)\sigma_j^2}{\sigma_{zj}^2} \tag{12}$$

follows the chi-square distribution with $m-1$ degree of freedom, m is the number of observations in each column and s is the seasonal lag.

The interval (12) with $100(1-\alpha)\%$ degree of confidence.

2.2 Levene's Test for Constant Variance

The Levene's test statistic for the null hypothesis

$$H_0 : \sigma_i^2 = \sigma_j^2$$

$H_1 : \sigma_i^2 \neq \sigma_j^2$ for at least one $i \neq j$ is defined as

$$W = \frac{(N-K) \sum_{i=1}^k N_i \left(\bar{z}_i - \bar{z}_{..} \right)^2}{(k-1) \sum_{i=1}^k \sum_{j=1}^{N_i} \left(z_{ij} - \bar{z}_i \right)^2} \tag{13}$$

where k is the number of different groups, N_i is the number of cases in the i th group, Y_{ij} is the value of the j th observation in the i th group .

z_{ij} may be defined as deviation of y_{ij} from the mean (\bar{y}_i) or from the median (\bar{y}_i). That is

$$z_{ij} = y_{ij} - \bar{y}_i \text{ or } y_{ij} - \bar{y}_i \tag{14}$$

$$\bar{z}_i = \frac{1}{N_i} \sum_{j=1}^{N_i} z_{ij} \text{ is the mean of the } z_{ij} \text{ for group } i \tag{15}$$

$$\bar{z}_{..} = \frac{1}{N} \sum_{i=1}^k \sum_{j=1}^{N_i} z_{ij} \text{ is mean of all } z_{ij}. \tag{16}$$

The test statistic W approximately follows the F-distribution with $k - 1$ and $N - K$ degree of freedom. To suit the Buys-Ballot procedure, the Levene's test statistic is modified with

$$N = ms, \quad k = s, \quad N_i = m \text{ as}$$

$$W = \frac{(ms - s)}{s - 1} \left[\frac{\sum_{j=1}^s m(\bar{z}_i - \bar{z}_{..})^2}{\sum_{j=1}^s \sum_{i=1}^m (z_{ij} - \bar{z}_{..j})^2} \right] \tag{17}$$

$$= \frac{s(m - 1)}{s - 1} \left[\frac{m \sum_{i=1}^s (\bar{z}_i - \bar{z}_{..})^2}{\sum_{i=1}^s \sum_{j=1}^m (z_{ij} - \bar{z}_{..j})^2} \right] \tag{18}$$

3 Empirical Examples

This section discusses some of the empirical examples to illustrate the validity of the Chi-Square test. The empirical examples consist of both simulated and real life data. Sections 3.1 and 3.2 contain simulation results for the mixed model and real life example respectively.

3.1 Simulations Results using the Mixed Model

The simulated series used consists 100 data set of 120 observations each simulated from

$$X_t = (a + bt) \times S_t + e_t, \text{ with } a = 1, \quad b = 0.02, \quad e_t \sim N(0,1). \text{ with } S_1 = 0.98, \quad S_2 = 0.80, \quad S_3 = 0.88,$$

$$S_4 = 1.04, \quad S_5 = 0.96, \quad S_6 = 1.22, \quad S_7 = 1.27, \quad S_8 = 1.32, \quad S_9 = 0.96, \quad S_{10} = 0.80, \quad S_{11} = 0.88, \quad S_{12} = 0.89$$

Using the MINITAB 17.0 version. The results of calculated value of the statistic from the simulated series are given in Table 2. The critical values at 5% level of significance and which for $m-1 = 9$ degree of freedom, equal to 2.7 and 19.0. Under the null hypothesis that the appropriate model is mixed, the calculated value of the statistic in (12) is expected to lie within the interval, otherwise, it will be concluded that the data does not admit mixed model. The calculated values of the statistic from the simulated series are listed in Table 2. When compared with the interval 2.7 and 19.0, the calculated values of the statistic lie within the interval in 100 out of the 100 stimulations. This shows that the test is capable of identifying the mixed model successfully 100 percent of the times. This expresses a high degree of confidence in the Chi-Square test.

Table 2. Calculated Chi-Square for Mixed Model: The critical values for $m - 1 = 9$ degree of freedom are 2.7 and 19

Col	Series									
	1	2	3	4	5	6	7	8	9	10
1	10.4821	5.0921	7.5963	7.7977	7.1774	9.3257	9.4257	3.1520	3.7203	9.0872
2	9.9157	10.8037	14.0504	3.8469	6.9436	5.4959	5.4413	12.9076	15.0742	12.7306
3	13.0416	5.1566	11.2798	8.6702	13.7272	8.7078	13.9401	6.8329	8.4794	6.9740
4	6.9085	12.2505	12.4492	14.8293	4.9423	7.2650	7.9545	7.3044	12.3286	12.3963
5	16.1748	10.4191	8.4801	10.5545	4.2798	17.1571	3.9333	5.2740	4.9656	12.9938
6	8.2899	16.4968	6.5897	3.3076	7.5989	9.7034	7.8557	15.1125	7.5076	6.6820
7	8.0266	10.3324	9.5820	16.5750	16.5760	4.9320	6.0845	9.1149	6.6597	12.0505
8	16.9952	2.5788	15.9770	9.7204	7.0275	13.3809	14.6644	9.7687	17.1803	9.2279
9	7.5575	3.3121	4.8592	7.8997	3.3922	4.2077	8.2510	4.6430	8.8565	8.9871
10	7.0924	14.5750	5.2015	8.1676	15.4410	11.3758	8.8120	13.1021	6.3318	3.9990
11	15.7355	12.9050	12.1345	6.8173	12.7856	4.7580	6.8402	6.0106	5.0302	7.8786
12	4.8516	8.6963	10.7143	2.7000	18.1478	5.0776	3.2408	12.7302	6.2707	8.8515
Decision	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept

Table 2. Calculated chi-square for mixed model

Col	Series									
	11	12	13	14	15	16	17	18	19	20
1	9.6094	14.2685	8.9455	4.3441	7.5303	13.2527	10.9166	5.9109	11.4996	4.7858
2	11.2284	10.0976	5.4781	12.1527	6.7144	8.8954	7.5931	7.3266	10.3937	6.3749
3	7.0312	5.5312	6.2797	15.7758	9.4296	5.5272	4.8595	10.4684	5.0037	2.0421
4	5.9895	9.5869	10.2811	9.0821	12.0272	8.1651	5.9885	11.6237	8.5224	7.9510
5	6.5379	4.7852	6.8220	7.7652	3.9335	6.2330	8.0709	4.5609	7.6014	13.8613
6	13.8824	5.7359	6.9559	6.2243	13.9755	9.0818	7.1403	13.1169	4.8691	5.9653
7	6.1482	7.9960	14.7312	13.7218	8.2492	11.2175	10.9352	16.8010	6.7297	8.7209
8	15.7446	5.4835	7.6453	6.5535	12.3519	5.6445	8.2929	5.1867	4.7716	17.7810
9	6.3454	6.6425	6.7641	13.7449	5.9355	8.0525	8.9194	4.1364	10.5184	7.6881
10	9.0031	7.5444	12.1063	5.8258	8.6373	11.5941	9.8175	15.5784	5.3195	4.6014
11	4.7313	16.6570	7.5044	3.1541	15.2454	6.3591	4.1399	4.2341	7.3711	10.6682
12	9.2021	7.9693	15.0332	9.3116	5.2108	6.6369	10.7285	10.6694	18.1081	7.5191
Decision	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept

Table 2. Calculated chi-square for mixed model

Col	Series									
	21	22	23	24	25	26	27	28	29	30
1	4.9925	6.6872	9.3019	8.4440	16.1231	4.6607	5.6206	12.8225	7.5324	7.2493
2	9.9520	7.6571	5.1852	3.8616	5.1936	16.0880	6.3393	9.3057	7.9772	11.2456
3	15.7739	9.2675	7.5613	7.3520	9.5426	3.5049	7.1022	7.4936	6.2753	13.7839
4	7.9491	8.8898	6.4351	3.4607	4.3826	8.9271	5.1558	6.6446	5.9960	4.3721
5	15.4095	7.8317	12.7285	2.6768	5.3428	8.9074	8.3408	14.2036	6.3303	5.2609
6	4.4069	11.7922	7.8317	16.1478	8.3204	10.8279	17.3949	7.7486	8.6958	6.9722

Col	Series									
	21	22	23	24	25	26	27	28	29	30
7	10.5531	10.0569	9.3096	14.4791	7.5278	4.7670	7.9232	10.5640	13.9507	8.6635
8	9.1184	5.5631	12.5454	7.8327	11.9395	9.5254	8.5931	9.3800	10.2167	8.2705
9	8.0141	8.2588	7.5039	6.1646	9.3666	5.1337	5.9292	8.7776	6.7719	7.2804
10	7.0412	3.1853	8.2615	11.9400	7.4869	5.5111	7.5531	13.3698	15.8669	10.7257
11	2.3845	8.1368	4.6635	3.9207	6.7324	14.2902	6.2990	8.0857	17.5143	10.1542
12	8.7773	15.5007	9.8291	3.6141	10.8442	10.7897	12.2908	10.2846	13.0388	10.5793
Decision	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept

Table 2. Calculated chi-square for mixed model

Col	Series									
	31	32	33	34	35	36	37	38	39	40
1	4.3301	5.2052	11.6108	8.4196	13.2948	8.7634	9.7346	7.9785	7.0880	5.2840
2	12.5402	8.0585	4.8702	8.7460	3.7832	3.6429	6.8156	7.5950	2.7919	18.6303
3	5.3371	6.7496	7.6270	5.8716	6.0977	4.2432	3.7521	3.4111	8.3039	3.9887
4	9.4842	14.6409	2.7816	6.4014	6.2389	2.0595	5.8052	5.1196	6.9820	3.5517
5	3.4534	8.2087	11.4106	4.0613	5.8798	6.6919	8.6991	2.1725	12.3841	10.2771
6	19.0150	8.7136	5.6496	14.7791	2.0007	17.3150	10.3676	17.1883	8.5542	5.1544
7	13.7678	17.1864	9.5216	11.3132	5.1484	7.9064	13.3484	6.0798	0.6810	4.6176
8	11.9285	9.2793	9.5300	10.9811	6.4705	8.3198	0.1218	4.7634	8.9523	7.3645
9	12.2726	7.0765	12.9150	11.2232	9.3524	6.5209	5.8827	10.9141	4.8175	14.1636
10	4.3265	9.1949	6.7857	10.5474	0.8313	17.9135	2.8196	8.8829	7.6546	6.6831
11	10.4583	3.4714	9.7626	5.9101	0.5449	6.0276	2.9715	12.7119	7.4249	4.9223
12	8.4716	4.9938	9.8488	5.7529	8.1471	4.4695	5.8214	3.4978	3.6641	7.1728
Decision	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept

Table 2. Calculated chi-square for mixed model

Col	Series									
	41	42	43	44	45	46	47	48	49	50
1	10.3691	8.5130	7.0330	4.3074	7.3176	6.8423	6.7538	9.7079	6.4727	7.0871
2	11.6339	7.8723	6.2565	12.1941	1.3870	3.4437	6.2822	0.9314	5.9619	7.611
3	4.4322	9.6867	14.3807	3.1714	9.9120	3.7780	0.9761	7.3828	0.6245	8.2764
4	8.5585	4.6870	4.5882	9.7055	2.9632	8.8073	4.5320	7.9623	5.9960	7.0251
5	5.6940	4.8856	10.7079	6.3863	0.2996	4.3244	5.4493	0.7636	12.7787	5.6149
6	7.5931	11.1882	5.7294	4.9469	9.5238	6.4484	9.4625	6.7796	5.4770	7.0564
7	15.1459	3.8121	0.5215	14.1749	1.2227	12.5596	8.8130	6.9940	8.6699	7.7016
8	6.3324	8.1706	2.8706	14.4109	4.4109	14.4099	7.0864	4.2041	1.1815	8.7890
9	7.8135	3.4254	9.6682	9.4238	8.8953	7.4470	6.5467	5.0319	5.9216	11.6033
10	3.0301	10.6712	6.3942	15.3995	7.6932	10.7740	4.9069	14.8724	3.6535	18.5999
11	11.0178	18.1801	8.6928	11.8529	16.1330	12.1989	12.4033	4.3272	5.8165	5.5348
12	12.5979	10.8027	7.3551	12.0797	6.6074	7.7126	8.1004	5.4474	8.1823	10.3745
Decision	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept

3.2 Real life example

The real life example is based on the monthly time series data on church marriages in Owerri, Imo State, Nigeria, for the period 2008 to 2019 given in Appendix A while the time plots are in Figs 1 and 2. The column (monthly) variances are shown in Table 3. The first step is to check whether the data admits the additive model. The modified Levene’s test statistic shown in (18) is used. The null hypothesis that the data admits additive model is reject if W is greater than the tabulated value, for which $\alpha = 0.05$ level of significance and $m - 1 = 11$ degree of freedom equal to 1.82 or do not reject H_0 otherwise. When compared with the critical value (1.82). W is greater, suggesting that the data does not admit the additive model.

Table 3. Deviations of the observed values from means ($Z_{ij} = \left| y_{ij} - \bar{y}_{.j} \right|$)

	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sept.	Oct.	Nov.	Dec.	total	$\bar{z}_{.j}$	$\sigma_{.j}$
2008	6.42	4.75	8.17	13.08	2.17	3.75	3.67	3.58	0.50	10.67	5.25	12.67	74.67	6.22	4.09
2009	6.42	1.75	1.83	2.08	5.17	0.75	0.67	3.58	0.50	2.33	6.75	6.67	38.50	3.21	2.42
2010	5.58	1.75	0.83	2.92	0.83	0.75	2.33	6.42	4.50	4.33	1.75	4.67	36.67	3.06	1.98
2011	3.42	0.25	4.83	8.08	3.83	1.25	1.33	4.42	3.50	5.33	2.25	2.33	40.83	3.40	2.14
2012	5.42	1.25	0.17	12.92	0.83	5.25	2.67	1.58	25.50	9.33	9.25	20.33	94.50	7.88	8.13
2013	2.58	0.75	1.17	16.92	5.17	6.25	0.33	12.58	5.50	16.67	16.25	3.33	87.50	7.29	6.50
2014	3.58	5.25	0.83	2.92	10.17	2.25	0.33	0.42	0.50	5.33	11.25	8.33	51.17	4.26	3.87
2015	7.58	5.25	1.17	4.92	3.17	2.75	0.33	1.58	2.50	1.33	5.75	5.33	41.67	3.47	2.25
2016	1.58	1.25	3.17	10.08	1.83	3.25	0.67	0.58	3.50	10.67	0.25	7.67	44.50	3.71	3.70
2017	1.58	2.25	0.17	3.08	7.83	4.75	3.33	2.42	5.50	13.33	10.75	0.33	55.33	4.61	4.12
2018	3.58	4.75	0.83	2.92	5.83	2.75	1.67	5.42	4.50	9.33	9.75	2.33	53.67	4.47	2.80
2019	4.42	1.75	4.83	7.08	4.83	2.75	1.33	4.42	4.50	12.67	9.75	10.67	69.00	5.75	3.58
Total	52.17	31.00	28.00	87.00	51.67	36.50	18.67	47.00	61.00	101.33	89.00	84.67	688.00		
$\bar{z}_{.j}$	4.35	2.58	2.33	7.25	4.31	3.04	1.56	3.92	5.08	8.44	7.42	7.06		4.78	
$\sigma_{.j}$	1.96	1.86	2.46	4.99	2.82	1.74	1.19	3.31	6.69	4.72	4.63	5.52			4.32

Table 4. Square of deviations of the observed values from seasonal means

	$\left(z_{ij} - \bar{z}_{.j} \right)^2$												
	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sept.	Oct.	Nov.	Dec.	total
2008	4.28	4.69	34.03	34.03	4.57	0.50	4.46	0.11	21.01	4.94	4.69	31.48	148.80
2009	4.28	0.69	0.25	26.69	0.74	5.25	0.79	0.11	21.01	37.35	0.44	0.15	97.76
2010	1.53	0.69	2.25	18.78	12.06	5.25	0.60	6.25	0.34	16.90	32.11	5.71	102.47
2011	0.87	5.44	6.25	0.69	0.22	3.21	0.05	0.25	2.51	9.68	26.69	22.30	78.17
2012	1.14	1.78	4.69	32.11	12.06	4.88	1.23	5.44	416.84	0.79	3.36	176.30	660.63
2013	3.11	3.36	1.36	93.44	0.74	10.29	1.49	75.11	0.17	67.60	78.03	13.85	348.58
2014	0.58	7.11	2.25	18.78	34.35	0.63	1.49	12.25	21.01	9.68	14.69	1.63	124.46
2015	10.47	7.11	1.36	5.44	1.30	0.09	1.49	5.44	6.67	50.57	2.78	2.97	95.69
2016	7.64	1.78	0.69	8.03	6.11	0.04	0.79	11.11	2.51	4.94	51.36	0.37	95.38
2017	7.64	0.11	4.69	17.36	12.45	2.92	3.16	2.25	0.17	23.90	11.11	45.19	130.95
2018	0.58	4.69	2.25	18.78	2.33	0.09	0.01	2.25	0.34	0.79	5.44	22.30	59.86
2019	0.005	0.69	6.25	0.03	0.28	0.09	0.05	0.25	0.34	17.83	5.44	13.04	44.29
Total	42.14	38.17	66.33	274.17	87.21	33.23	15.63	120.83	492.92	244.96	236.17	335.30	1987.05

Table 5. Calculation of $m \left(\bar{z}_{.j} - \bar{z}_{..} \right)^2$ m = 12

$\bar{z}_{.j}$	$\bar{z}_{..}$	$\bar{z}_{.j} - \bar{z}_{..}$	$\left(\bar{z}_{.j} - \bar{z}_{..} \right)^2$	$12 \times \left(\bar{z}_{.j} - \bar{z}_{..} \right)^2$
4.35	4.78	-0.43	0.18	2.22
2.58	4.78	-2.20	4.84	58.08
2.33	4.78	-2.45	6.00	72.03
7.25	4.78	2.47	6.10	73.21
4.31	4.78	-0.47	0.22	2.65
3.04	4.78	-1.74	3.03	36.33
1.56	4.78	-3.22	10.37	124.42
3.92	4.78	-0.86	0.74	8.88
5.08	4.78	0.30	0.09	1.08
8.44	4.78	3.66	13.40	160.75
7.42	4.78	2.64	6.97	83.64
7.06	4.78	2.28	5.20	62.38
				685.66

From Appendix A and Table 5

$$W = \frac{12 * (12 - 1) (685.66)}{(12 - 1) (1987.05)} = \frac{90507.12}{21857.55} = 4.14$$

Having admitted that the data does not additive model, the choice now choose lies between mixed and multiplicative models. According to the chi-square test by Nwogu, et al. [5] and Dozie et al. [6] the null hypothesis that the data admits the mixed model is rejected, if the statistic defined in (12) lies outside the

interval $\left[\chi_{\frac{\alpha}{2}, (m-1)}^2, \chi_{1-\frac{\alpha}{2}, (m-1)}^2 \right]$ which for $\alpha = 0.05$ level of significance and $m - 1 = 11$ degrees of freedom,

equals (3.8, 21.9) or do not reject H_0 otherwise, and from (12) the calculated values, χ_{cal}^2 given in Table 6 were obtained. When compared with the critical values (3.8 and 21.9), all the calculated values lie outside the interval, indicating that the data does not admit mixed model.

However, there is indication the choice of model may be affected by violation of the underlying assumptions, therefore, there is need to evaluate data for transformation to meet the constant variance and normality assumptions in the distribution. When the column variances of the transformed series given in Table 7 are subjected to test for constant variance, the calculated Levene's test statistic (0.78) is less than the tabulated (1.82) at $\alpha = 0.05$ level of significance and $m - 1 = 11$ degrees of freedom. This indicates that the variance is constant and the transformed series admits additive model.

Table 6. Seasonal effects (S_j), estimate of the column variance ($\hat{\sigma}_j^2$) and Calculated Chi-square (χ_{cal}^2)

j	1	2	3	4	5	6	7	8	9	10	11	12
S_j	1.68	0.77	0.77	2.18	1.15	0.54	0.30	0.67	0.63	0.98	0.98	1.63
$\hat{\sigma}_j^2$	24.5	10.8	11.9	82.3	28.2	13.1	4.1	27.7	73.0	100	81.5	84.8
χ_{cal}^2	0.03	0.06	0.06	0.05	0.07	0.14	0.14	0.19	0.58	0.32	0.26	0.10

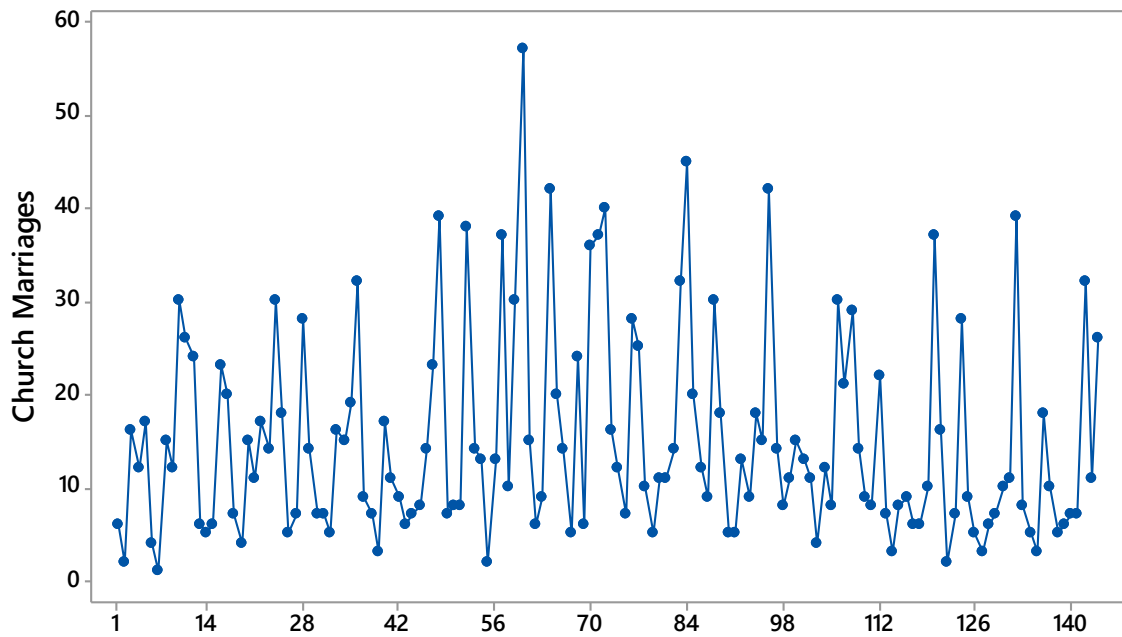


Fig. 1. Time plot of original series of church marriages

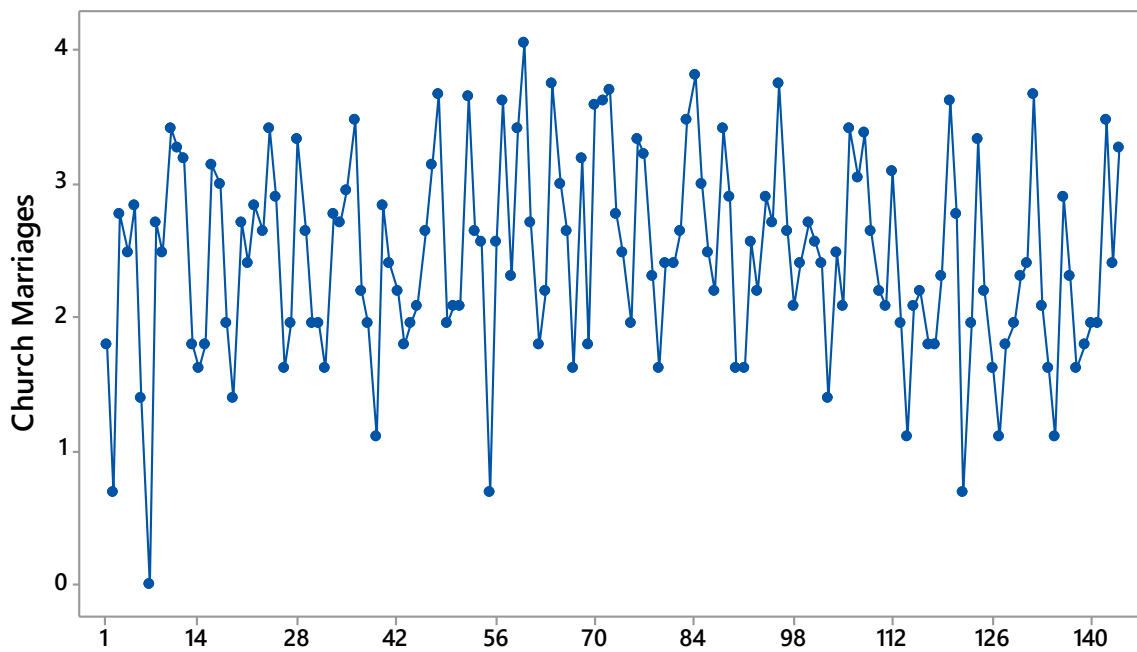


Fig. 2. Time plot of transformed series of church marriages

From appendix B and Table 6, $\sigma_1^2 = 1$, $b = 0.1143$, $n = 144$, $m = 12$

Therefore, from (6), $\sigma_{z_j}^2 = (0.1143)^2 \times 144 \left(\frac{144+12}{12} \right) S_j^2 + 1$

Table 7. Transformed series of Deviations of the Observed Values from Means ($Z_{ij} = \left| y_{ij} - \bar{y}_{.j} \right|$)

	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sept.	Oct.	Nov.	Dec.	total	$\bar{z}_{.j}$	$\sigma_{.j}$
2008	0.64	1.08	0.81	0.67	0.20	0.56	1.42	0.37	0.19	0.57	0.32	0.40	7.22	0.60	0.36
2009	0.64	0.16	0.17	0.02	0.36	0.00	0.03	0.37	0.11	0.00	0.30	0.17	2.35	0.20	0.19
2010	0.46	0.16	0.02	0.17	0.00	0.00	0.53	0.73	0.48	0.12	0.00	0.11	2.79	0.23	0.25
2011	0.24	0.17	0.86	0.33	0.24	0.25	0.38	0.40	0.21	0.19	0.19	0.09	3.55	0.30	0.20
2012	0.49	0.31	0.12	0.48	0.00	0.62	0.72	0.22	1.32	0.53	0.46	0.47	5.74	0.48	0.34
2013	0.27	0.02	0.23	0.58	0.36	0.70	0.19	0.84	0.50	0.75	0.67	0.12	5.23	0.44	0.27
2014	0.34	0.71	0.02	0.17	0.58	0.36	0.19	0.06	0.11	0.19	0.52	0.23	3.49	0.29	0.22
2015	0.56	0.71	0.23	0.24	0.26	0.33	0.19	0.22	0.09	0.06	0.23	0.16	3.31	0.28	0.19
2016	0.20	0.31	0.44	0.45	0.07	0.46	0.03	0.14	0.21	0.57	0.10	0.21	3.19	0.27	0.18
2017	0.20	0.42	0.12	0.07	0.69	0.84	0.66	0.14	0.50	1.04	0.64	0.04	5.37	0.45	0.33
2018	0.34	1.08	0.02	0.17	0.44	0.33	0.32	0.55	0.35	0.53	0.54	0.09	4.75	0.40	0.28
2019	0.36	0.16	0.86	0.27	0.33	0.33	0.38	0.40	0.35	0.64	0.54	0.32	4.93	0.41	0.19
Total	4.74	5.30	3.90	3.63	3.53	4.80	5.04	4.43	4.42	5.19	4.53	2.40	51.91		
$\bar{z}_{.j}$	0.40	0.44	0.32	0.30	0.29	0.40	0.42	0.37	0.37	0.43	0.38	0.20		0.36	
$\sigma_{.j}$	0.16	0.37	0.34	0.20	0.21	0.25	0.38	0.24	0.34	0.32	0.22	0.13			0.27

Table 8. Transformed Series of square of deviations of the observed values from seasonal means

	$\left(z_{ij} - \bar{z}_{.j} \right)^2$												
	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sept.	Oct.	Nov.	Dec.	total
2008	0.06	0.41	0.24	0.14	0.01	0.02	0.99	0.00	0.03	0.02	0.00	0.04	1.96
2009	0.06	0.08	0.02	0.08	0.00	0.16	0.15	0.00	0.07	0.18	0.01	0.00	0.81
2010	0.00	0.08	0.09	0.02	0.08	0.16	0.01	0.13	0.01	0.10	0.14	0.01	0.84
2011	0.02	0.07	0.29	0.00	0.00	0.02	0.00	0.00	0.02	0.06	0.03	0.01	0.54
2012	0.01	0.02	0.04	0.03	0.08	0.05	0.09	0.02	0.91	0.01	0.01	0.07	1.34
2013	0.01	0.18	0.01	0.08	0.00	0.09	0.05	0.22	0.02	0.10	0.08	0.01	0.85
2014	0.00	0.07	0.09	0.02	0.08	0.00	0.05	0.10	0.07	0.06	0.02	0.00	0.57
2015	0.03	0.07	0.01	0.00	0.00	0.00	0.05	0.02	0.07	0.14	0.02	0.00	0.43
2016	0.04	0.02	0.01	0.02	0.05	0.00	0.15	0.05	0.02	0.02	0.08	0.00	0.47
2017	0.04	0.00	0.04	0.05	0.16	0.20	0.06	0.05	0.02	0.37	0.07	0.03	1.08
2018	0.00	0.41	0.09	0.02	0.02	0.00	0.01	0.03	0.00	0.01	0.03	0.01	0.64
2019	0.002	0.08	0.29	0.00	0.00	0.00	0.00	0.00	0.00	0.04	0.03	0.01	0.46
Total	0.28	1.48	1.24	0.46	0.50	0.71	1.63	0.63	1.24	1.10	0.52	0.19	9.99

Table 9. Calculation of $m \left(\bar{z}_{.j} - \bar{z}_{..} \right)^2$ m=12

$\bar{z}_{.j}$	$\bar{z}_{..}$	$\bar{z}_{.j} - \bar{z}_{..}$	$\left(\bar{z}_{.j} - \bar{z}_{..} \right)^2$	$12 \times \left(\bar{z}_{.j} - \bar{z}_{..} \right)^2$
0.4	0.36	0.040	0.002	0.019
0.44	0.36	0.080	0.006	0.077
0.32	0.36	-0.040	0.002	0.019
0.3	0.36	-0.060	0.004	0.043
0.29	0.36	-0.070	0.005	0.059
0.4	0.36	0.040	0.002	0.019
0.42	0.36	0.060	0.004	0.043
0.37	0.36	0.010	0.000	0.001
0.37	0.36	0.010	0.000	0.001
0.43	0.36	0.070	0.005	0.059
0.38	0.36	0.020	0.000	0.005
0.2	0.36	-0.160	0.026	0.307
				0.653

From Appendix B and Table 9

$$W = \frac{12 * (12 - 1) (0.653)}{(12 - 1) (9.99)} = \frac{86.196}{109.89} = 0.78$$

4 Concluding Remarks

This article has discussed the Chi-Square test in time series data. The Chi-Square test is based on Chi-Square distribution. Although time series data does not satisfy all the assumptions of most common statistical tests, the Chi-Square test appears to be the most efficient among them. The Chi-Square test is able to identify the mixed model successfully in one hundred (100) out of the one hundred (100) simulations. Secondly, the transformed series shown in Appendix B and Table 9 admits additive model. This further confirms that the appropriate model of original data given in Appendix A and Table 5 is multiplicative. In view of this, it is recommended

that a study data should be evaluated for assumptions of time series model before applying test for choice of model.

Competing Interests

Authors have declared that no competing interests exist.

References

- [1] Kendal MG, Ord JK. Time Series 3rd ed. Charles Griffin, London; 1990.
- [2] Chatfield C. The analysis of time Series: An introduction. Chapman and Hall,/CRC Press, Boca Raton; 2004.
- [3] Puerto J, Rivera MP. Descriptive analysis of time series applied to housing prices in Spain, Management Mathematics for European Schools 94342-CP-2001- DE – COMENIUS – C21; 2001.
- [4] Iwueze IS, Nwogu EC, Ohakwe J, Ajaraogu JC. Uses of the Buys-Ballot table in time series analysis. Applied Mathematics Journal. 2011;2:633-645.
- [5] Nwogu EC, Iwueze IS, Dozie KCN, Mbachu HI. Choice between mixed and multiplicative models in time series decomposition. International Journal of Statistics and Applications. 2019;9(5):153-159.
- [6] Dozie KCN, Mbachu HI, Raymond MC. The analysis of mixed model in time series decomposition. International Journal of Probability and Statistics. 2020;9(1):1-6.
- [7] Akpanta AC, Iwueze IS. On applying the Bartlett transformation method to time series data. Journal of Mathematical Sciences. 2009;20(5):227-243.
- [8] Dozie KCN, Ijomah MA. A comparative study on additive and mixed models in descriptive time series. American Journal of Mathematical and Computer Modelling. 2020;5(1):12-17.
- [9] Dozie KCN, Nwanya JC. Comparison of mixed and multiplicative models, when trend cycle components is linear. Asian Journal of Advance Research and Reports. 2020;12(4):32-42.
- [10] Dozie KCN. Buys-Ballot estimates for mixed model in descriptive time series. International Journal of Theoretical and Mathematical Physics. 2020;10(1):22-27.
- [11] Dozie KCN, Uwaezuoke MU. Procedure for estimation of additive time series model. International Journal of Research and Scientific Innovation. 2021;8(2):251-256.
- [12] Dozie KCN, Ibebuogu CC. Road traffic offences in Nigeria: An empirical analysis using buys-ballot approach. Asian Journal of Probability and Statistics. 2021;12(1):68-78.
- [13] Dozie KCN, Ihekuna SO. Additive seasonality in time series using row and overall sample variances of the buys-ballot table. Asian Journal of Probability and Statistics. 2022;18(3):1-9.

Appendix A. Actual series of church marriages in Imo State, Nigeria (2008 – 2019)

	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sept.	Oct.	Nov.	Dec.	total	\bar{x}_i	σ_i
2008	6	2	16	12	17	4	1	15	12	30	26	24	165	13.75	9.53
2009	6	5	6	23	20	7	4	15	11	17	14	30	152	13.82	8.28
2010	18	5	7	28	14	7	7	5	16	15	19	32	166	15.09	9.02
2011	9	7	3	17	11	9	6	7	8	14	23	39	153	12.75	9.88
2012	7	8	8	38	14	13	2	13	37	10	30	57	229	20.82	17.16
2013	15	6	9	42	20	14	5	24	6	36	37	40	245	22.27	14.37
2014	16	12	7	28	25	10	5	11	11	14	32	45	216	18.00	11.99
2015	20	12	9	30	18	5	5	13	9	18	15	42	187	17.00	10.96
2016	14	8	11	15	13	11	4	12	8	30	21	29	176	14.67	8.09
2017	14	9	8	22	7	3	8	9	6	6	10	37	131	11.91	9.72
2018	16	2	7	28	9	5	3	6	7	10	11	39	136	12.36	11.42
2019	8	5	3	18	10	5	6	7	7	32	11	26	138	11.50	9.12
Total	149	81	40	301	178	93	56	137	138	232	249	440			
-	12.42	6.75	8.00	25.08	14.83	7.75	4.67	11.42	11.50	19.33	20.75	36.67			
$y_{.j}$															
$\sigma_{.j}$	4.94	3.28	5.57	9.07	5.31	3.62	2.02	5.25	8.54	10.00	9.03	9.21			

Appendix B. Transformed series of church marriages in Imo State, Nigeria (2008 – 2019)

	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	T_i	\bar{X}_i	$\hat{\sigma}_i$
2008	1.7918	0.6931	2.7726	2.4849	2.8332	1.3863	0.0000	2.7081	2.4849	3.4012	3.2581	3.1781	26.9922	2.2494	1.0693
2009	1.7918	1.6094	1.7918	3.1355	2.9957	1.9459	1.3863	2.7081	2.3979	2.8332	2.6391	3.4012	28.6358	2.3863	0.6629
2010	2.8904	1.6094	1.9459	3.3322	2.6391	1.9459	1.9459	1.6094	2.7726	2.7081	2.9444	3.4657	29.8091	2.4841	0.6478
2011	2.1972	1.9459	1.0986	2.8332	2.3979	2.1972	1.7918	1.9459	2.0794	2.6391	3.1355	3.6636	27.9253	2.3271	0.6728
2012	1.9459	2.0794	2.0794	3.6376	2.6391	2.5649	0.6931	2.5649	3.6109	2.3026	3.4012	4.0431	31.5622	2.6302	0.9298
2013	2.7081	1.7918	2.1972	3.7377	2.9957	2.6391	1.6094	3.1781	1.7918	3.5835	3.6109	3.6889	33.5321	2.7943	0.796
2014	2.7726	2.4849	1.9459	3.3322	3.2189	2.3026	1.6094	2.3979	2.3979	2.6391	3.4657	3.8067	32.3738	2.6978	0.649
2015	2.9957	2.4849	2.1972	3.4012	2.8904	1.6094	1.6094	2.5649	2.1972	2.8904	2.7081	3.7377	31.2866	2.6072	0.6448
2016	2.6391	2.0794	2.3979	2.7081	2.5649	2.3979	1.3863	2.4849	2.0794	3.4012	3.0445	3.3673	30.5509	2.5459	0.5655
2017	2.6391	2.1972	2.0794	3.091	1.9459	1.0986	2.0794	2.1972	1.7918	1.7918	2.3026	3.6109	26.825	2.2354	0.6476
2018	2.7726	0.6931	1.9459	3.3322	2.1972	1.6094	1.0986	1.7918	1.9459	2.3026	2.3979	3.6636	25.7508	2.1459	0.8472
2019	2.0794	1.6094	1.0986	2.8904	2.3026	1.6094	1.7918	1.9459	1.9459	3.4657	2.3979	3.2581	26.3952	2.1996	0.7051
$T_{.j}$	29.224	21.278	23.551	37.916	31.621	23.307	17.002	28.097	27.496	33.958	35.306	42.885	351.639		
-	2.4353	1.7732	1.9625	3.1597	2.6351	1.9422	1.4168	2.3414	2.2913	2.8299	2.9422	3.5737		2.4419	
$\bar{X}_{.j}$															
$\hat{\sigma}_{.j}$	0.4428	0.5896	0.4773	0.3758	0.3746	0.4889	0.5835	0.4537	0.5108	0.5519	0.45	0.2472			0.7493

© 2022 Dozie and Ihekuna; This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Peer-review history:

The peer review history for this paper can be accessed here (Please copy paste the total link in your browser address bar)

<https://www.sdiarticle5.com/review-history/90872>