Hacettepe Journal of Mathematics and Statistics
Volume 45 (6) (2016), 1743 – 1754

Linear contrasts in one-way classification AR(1) model with gamma innovations

Birdal Şenoğlu*† and Özlem Türker Bayrak‡

Abstract

In this study, the explicit estimators of the model parameters in one-way classification AR(1) model with gamma innovations are derived by using modified maximum likelihood (MML) methodology. We also propose a new test statistic for testing linear contrasts. Monte Carlo simulation results show that the MML estimators have higher efficiencies than the traditional least squares (LS) estimators and the proposed test has much better power and robustness properties than the normal-theory test.

Keywords: Autoregressive model, linear contrasts, nonnormality, robustness, modified likelihood, gamma distribution.

2000 AMS Classification: AMS 62F03, 62F10

Received : 11.12.2012 *Accepted* : 16.02.2016 *Doi* : 10.15672/HJMS.20164515996

1. Introduction

Linear contrasts are widely used to make comparisons among the treatment means of interest. The usage of them require the independence assumption for the observations in each treatment. However, in numerous situations, the present state of a variable in each treatment is influenced by its past and this gives rise to autocorrelated time series structure. For instance in the agricultural and the biological sciences, the observations that are recorded over some time-space coordinate are extremely common, see, for example [7]. Some of the reasons for the lack of independence are (see [15]):

- Biased measurements,
- A poor allocation of treatments to experimental units,
- Adjacent experimental units or plots in a field.

*Department of Statistics, Science Faculty, Ankara University, Dö Gol Caddesi, 06100, Tandoğan, Ankara, Turkey, Email: senoglu@science.ankara.edu.tr

†Corresponding Author.

‡Statistics Unit, Cankaya University, Ankara, Turkey, Email: ozlemt@cankaya.edu.tr

Another standard assumption is that the error terms are i.i.d (identically and independently distributed) as normal $N(\mu, \sigma^2)$. From the practical point of view, this assumption is also not realistic since nonnormal error distributions are more prevalent. There exists huge literature on the subject of nonnormal error distributions, see, for example, [8], [11], [5], [20], [31].

The normal theory test statistics for testing linear contrasts have low efficiencies when the normality assumption is not satisfied, see [18]. However, they can still be used for the situations where the normality assumption is violated to a slight or moderate degree. On the other hand, if the independence assumption is not met, traditional test statistics do not work well and give misleading results even if the observations exhibit low levels of correlation over time, see [16] and [12].

In recent years, the MML method has been applied to various time series models by Tiku and his colleagues. [21] developed a unit root test for the AR(1) model. The first order autoregressive model, AR(1), has been considered in [22] with asymmetric innovations of the gamma type. [24] extended the results of [22] to the symmetric non-normal innovations. [25] gave some engineering applications of the AR(1) models with nonnormal errors. [23] and [1] considered the simple regression model with first-order autoregressive errors when the error distribution is symmetric and asymmetric nonnormal, respectively. [26] and [3] extended this methodology to various independent sources of information and to multiple autoregressive model under non-normality; respectively. [31] extended the results of [23] to the generalized logistic distribution family representing very wide skew distributions ranging from highly right skewed to the highly left skewed.

Skew distributions are observed frequently in the context of experimental design; see for example, [18] and [17]. In their real life applications, they observed that the error terms are distributed as Generalized Logistic(b,σ) with shape parameters $b = 1, 2, 6$ and Weibull(p,σ) with shape parameter $p = 4$; respectively. Thus, positively skewed distributions fitted very well to the error terms. Therefore, different than the earlier studies, we assume that the error terms have Gamma which is another widely used and well known positive skewed distribution. Besides, we assume that the observations in each treatment are first order autocorrelated. This is the first study, dealing with both autocorrelation and non-normality in experimental design as far as we know. Thus, we aim to fill this gap in the literature.

We derive the estimators of the model parameters in this one-way classification model by using MML methodology. The methodology was first initiated by [19]. We also propose a new test statistic based on these MML estimators for testing linear contrasts and show that our solutions are much more efficient than the traditional normal-theory solutions.

The methodology developed in this paper can be extended to other designs, time series models (e.g. factorial designs AR(2) model) and any location-scale distribution (e.g., long-tailed symmetric and short-tailed symmetric distributions).

2. One-way classification AR(1) model

Consider the following one-way classification model with first-order autoregressive errors:

$$y_{i,j} - \phi y_{i,j-1} = \mu_i + e_{i,j}, \qquad -1 < \phi < 1;\ -\infty < \mu_i < \infty; \quad i = 1, \ldots, a;\ j = 1, \ldots, n \tag{2.1}$$

or alternatively reparametrized as

$$y_{i,j} - \phi y_{i,j-1} = \mu + \tau_i + e_{i,j}, \qquad -1 < \phi < 1; -\infty < \tau_i < \infty; \tag{2.2}$$
$$-\infty < \mu < \infty; i = 1, \dots, a; j = 1, \dots, n$$

where $y_{i,j}$ is the jth observation in the ith treatment; μ is the constant representing the overall mean; μ_i is the mean of the ith treatment; τ_i is the ith treatment effect and $e_{i,j}$ is the error term.

Without loss of generality, we assume that $\sum \tau_i = 0$. Besides assume that $e_{i,j}$ are iid and have the gamma distribution

$$f(e) = \frac{1}{\sigma^k \Gamma(k)} exp(-\frac{e}{\sigma}) e^{k-1}; \qquad 0 < e < \infty \tag{2.3}$$

where k is the shape parameter and is assumed to be known. Conditional on $y_{i,0}$, the likelihood function ignoring the constant term which has no effect on the estimators is

$$L = \frac{1}{\sigma^n} e^{\sum_{i=1}^{a} \sum_{j=1}^{n} z_{i,j}} \prod_{i=1}^{a} \prod_{j=1}^{n} z_{i,j}^{k-1} \tag{2.4}$$

where $z_{i,j} = e_{i,j}/\sigma = (y_{i,j} - \phi y_{i,j-1} - \mu - \tau_i)/\sigma$.

The corresponding likelihood equations can be written as

$$\frac{\partial lnL}{\partial \mu} = \frac{N}{\sigma} - \frac{(k-1)}{\sigma} \sum_{i=1}^{a} \sum_{j=1}^{n} g(z_{i,j}) = 0$$

$$\frac{\partial lnL}{\partial \tau_i} = \frac{n}{\sigma} - \frac{(k-1)}{\sigma} \sum_{j=1}^{n} g(z_{i,j}) = 0$$

$$\frac{\partial lnL}{\partial \phi} = \frac{1}{\sigma} \sum_{i=1}^{a} \sum_{j=1}^{n} y_{i,j-1} - \frac{(k-1)}{\sigma} \sum_{i=1}^{a} \sum_{j=1}^{n} y_{i,j-1} g(z_{i,j}) = 0$$

$$\frac{\partial lnL}{\partial \sigma} = -\frac{N}{\sigma} + \frac{1}{\sigma} \sum_{i=1}^{a} \sum_{j=1}^{n} z_{i,j} - \frac{(k-1)}{\sigma} \sum_{i=1}^{a} \sum_{j=1}^{n} z_{i,j} g(z_{i,j}) = 0 \tag{2.5}$$

where $g(z) = 1/z$ and $N = an$: total number of observations.

These equations are in terms of $1/z_{i,j}$ and have no explicit solutions. Therefore they have to be solved by iteration which might be problematic especially when the data contains outliers, see, for example, [14], [27] and [28]. We, therefore, utilize the method of modified likelihood estimation which captures the beauty of maximum likelihood but alleviates its computational difficulties, see [20].

3. The MML estimators

The first step of obtaining the MML estimators is to express the likelihood equations (2.5) in terms of ordered $z_{i,(j)}$'s ($i = 1, \dots, a;\ j = 1, \dots, n$), since the complete sums are invariant to ordering. The second step is to linearize the term $g(z_{i,(j)}) = 1/z_{i,(j)}$ around $t_{(j)}$ by the use of the first two terms of a Taylor series expansion, since for large n, $z_{i,(j)}$ is close to its expected value $t_{(j)} = E(z_{i,(j)})$. Thus,

$$g(z_{i,(j)}) \cong g(t_{(j)}) + (z_{i,(j)} - t_{(j)}) \left\{ \frac{\partial g(z)}{\partial z} \right\}_{z=t_{(j)}} = \alpha_j - \beta_j z_{i,(j)} \tag{3.1}$$

where $\alpha_j = 2/t_{(j)}$ and $\beta_j = 1/t_{(j)}^2$. Although the exact values of the $t_{(j)}$ are available, for convenience, we use their approximate values generated from the equation $\frac{1}{\Gamma(k)}\int_0^{t_{(j)}} e^{-z}z^{k-1}dz = \frac{j}{n+1}$, $1 \le j \le n$ for each treatment (i.e., for $i = 1, \ldots, a$).

Incorporating the linear approximation 3.1 into the likelihood equations 2.5 yields the modified likelihood equations. Then the MML estimators are obtained by solving these modified likelihood equations as:

$$\widehat{\mu} = \widehat{\mu}_{.[.]} + \frac{\Delta}{m}\widehat{\sigma}, \quad \widehat{\tau}_i = \widehat{\mu}_{i[.]} - \widehat{\mu}_{.[.]},$$

$$\widehat{\phi} = K + D\widehat{\sigma}, \quad \widehat{\sigma} = \frac{B + \sqrt{B^2 + 4NC}}{2\sqrt{N(N-a-1)}} \tag{3.2}$$

where

$$\widehat{\mu}_{i[.]} = \frac{\sum_{j=1}^{n}\beta_j(y_{i,[j]} - \phi y_{i,[j]-1})}{m}, \ \widehat{\mu}_{.[.]} = \frac{\sum_{i=1}^{a}\sum_{j=1}^{n}\beta_j(y_{i,[j]} - \phi y_{i,[j]-1})}{am},$$

$$\Delta_j = \frac{1}{k-1} - \alpha_j, \quad \Delta = \sum_{j=1}^{n}\Delta_j, \quad m = \sum_{j=1}^{n}\beta_j,$$

$$K = \frac{\sum_{i=1}^{a}\sum_{j=1}^{n}\beta_j y_{i,[j]}y_{i,[j]-1} - \frac{1}{m}\sum_{i=1}^{a}(\sum_{j=1}^{n}\beta_j y_{i,[j]})(\sum_{j=1}^{n}\beta_j y_{i,[j]-1})}{\sum_{i=1}^{a}\sum_{j=1}^{n}\beta_j y_{i,[j]-1}^2 - \frac{1}{m}\sum_{i=1}^{a}(\sum_{j=1}^{n}\beta_j y_{i,[j]-1})^2},$$

$$D = \frac{\sum_{i=1}^{a}\sum_{j=1}^{n}(\Delta_j - \beta_j\frac{\Delta}{m})y_{i,[j]-1}}{\sum_{i=1}^{a}\sum_{j=1}^{n}\beta_j y_{i,[j]-1}^2 - \frac{1}{m}\sum_{i=1}^{a}(\sum_{j=1}^{n}\beta_j y_{i,[j]-1})^2},$$

$$B = (k-1)\sum_{i=1}^{a}\sum_{j=1}^{n}(y_{i,[j]} - \phi y_{i,[j]-1} - \widehat{\mu}_{i[.]})\Delta_j, \ \textit{and}$$

$$C = (k-1)\sum_{i=1}^{a}\sum_{j=1}^{n}\beta_j(y_{i,[j]} - \phi y_{i,[j]-1} - \widehat{\mu}_{i[.]})^2. \tag{3.3}$$

It is clear that the MML estimators have closed forms. It should also be noted that they have exactly the same forms as other MML estimators irrespective of the underlying distribution besides having the invariance property, see [20]. The MML estimators are known to be asymptotically fully efficient, i.e. they are unbiased and minimum variance bounds (MVB) estimators, see [4] and [29]. For small sample sizes, they have very little or no bias and the true variances of the MML estimators are very close to minimum variance bounds, see [28].

For the computation of the MML estimators $\widehat{\mu}$, $\widehat{\tau}_i$, $\widehat{\phi}$ and $\widehat{\sigma}$, first the ordered variates of $z_{i,j} = e_{i,j}/\sigma = (y_{i,j} - \phi y_{i,j-1} - \mu - \tau_i)/\sigma$ $(i = 1, \ldots, a;\ j = 1, \ldots, n)$ has to be obtained. Since the ordering of $z_{i,j}$ only depends on ϕ (μ and τ_i are additive constants and σ is positive), it is done by using the LS estimate $\widehat{\phi}_{LS}$ of ϕ as an initial estimate. Then using the concomitants $(y_{i,[j]}, y_{i,[j]-1})$ corresponding to ordered variates $w_{i,(j)} = y_{i,[j]} - \widehat{\phi}_{LS}y_{i,[j]-1}$, the MML estimates $\widehat{\mu}$, $\widehat{\tau}_i$, $\widehat{\phi}$ and $\widehat{\sigma}$ are calculated from 3.2. A second iteration is carried out by replacing $\widehat{\phi}_{LS}$ with $\widehat{\phi}$ in the ordering of $w_{i,(j)}$ variates and new $\widehat{\mu}$, $\widehat{\tau}_i$, $\widehat{\phi}$ and $\widehat{\sigma}$ values are calculated. This is repeated till the estimates stabilize sufficiently enough. In our computations, two iterations were enough. Actually, in literature based on MML, it can be seen that at most three iterations are enough.

4. Efficiency of the MML estimators

In practice the LS estimators are widely used which will be shown that they are considerably less efficient than the MML estimators. Relative efficiencies (RE) of the LS estimators defined as

$$RE = 100 \times (variance\,of\,MMLE)/(variance\,of\,LSE) \tag{4.1}$$

are calculated by simulation based on $[100000/n]$ Monte Carlo runs. Although much other values are tried, the simulation results performed for sample sizes $n = 30$, 60 and 120 with the shape parameter taking the values $k = 2$, 3, 5 and 10 for $\phi = 0.0$, 0.5 and 0.9 are given in Table 1. It must be noted that the values for other ϕ values including negative ones yield the similar results so that they are not reported.

The model parameters μ_i, τ_i and σ are set as 0, 0 and 1 without loss of generality. Realize that for $\phi = 0.0$, the model 2.1 turns to be the usual one-way classification where the errors are distributed as gamma rather than normal. In fact, this is by its own a contribution since the model parameters in one-way classification model have not been estimated with gamma distributions so far.

The LS estimators of the model parameters are given by

$$\widetilde{\mu}_i = \frac{\sum_{j=1}^{n}(y_{i,j} - \phi y_{i,j-1})}{n} - k\widetilde{\sigma}, \quad \widetilde{\mu} = \frac{\sum_{i=1}^{a}\sum_{j=1}^{n}(y_{i,j} - \phi y_{i,j-1}}{an} - k\widetilde{\sigma},$$

$$\widetilde{\tau}_i = \widetilde{\mu}_i - \widetilde{\mu}, \widetilde{\phi} = \frac{\sum_{i=1}^{a}\sum_{j=1}^{n} y_{i,j}y_{i,j-1} - \frac{1}{n}\sum_{i=1}^{a}(\sum_{j=1}^{n} y_{i,j})(\sum_{j=1}^{n} y_{i,j-1})}{\sum_{i=1}^{a}\sum_{j=1}^{n} y_{i,j-1}^2 - \frac{1}{n}\sum_{i=1}^{a}(\sum_{j=1}^{n} y_{i,j-1})^2},$$

$$\widetilde{\sigma}^2 = \frac{\sum_{i=1}^{a}\sum_{j=1}^{n}((y_{i,j} - \phi y_{i,j-1}) - \widetilde{\mu}_i)^2}{(N-a-1)k}. \tag{4.2}$$

Note that the LS estimators $\widetilde{\mu}$ and $\widetilde{\sigma}^2$ are corrected for bias so that they become comparable with MML estimators. Besides, the initial values $y_{i,0}$ are taken as $e_{i,0}/\sqrt{1-\phi^2}$, which is, in fact, Model II of [30].

It can be seen from Table 1 that the MML estimators are more efficient than the LS estimators especially for the small values of the shape parameter k. It should be noted that the relative efficiency of the LS estimators decrease as the sample size n increase. This is another result of interest.

Table 1. Simulated means (1), $n\times$variances (2) and the relative efficiencies (RE) of the LS and MML estimators.

n		$\widetilde{\mu}_i$	$\widehat{\mu}_i$	RE	$\widetilde{\tau}_i$	$\widehat{\tau}_i$	RE	$\widetilde{\phi}$	$\widehat{\phi}$	RE	$\widetilde{\sigma}$	$\widehat{\sigma}$	RE
							$k=2.0,\ \phi=0.0$						
30	(1)	0.073	0.160	30	0.001	0.001	28	-0.031	-0.009	39	0.991	0.973	53
	(2)	0.125	0.038		1.444	0.404		0.309	0.121		0.426	0.227	
60	(1)	0.031	0.087	24	0.000	0.000	25	-0.015	-0.002	30	0.997	0.982	51
	(2)	0.060	0.015		1.346	0.334		0.317	0.096		0.408	0.210	
120	(1)	0.019	0.050	19	-0.001	-0.001	21	-0.008	0.000	21	1.000	0.988	45
	(2)	0.032	0.006		1.289	0.269		0.344	0.070		0.424	0.193	
							$k=2.0,\ \phi=0.5$						
30	(1)	0.242	0.229	32	0.008	0.004	27	0.441	0.478	38	0.993	0.972	52
	(2)	0.229	0.073		1.710	0.455		0.262	0.099		0.417	0.215	
60	(1)	0.109	0.109	26	0.004	0.002	22	0.473	0.493	30	0.999	0.983	50
	(2)	0.105	0.027		1.500	0.326		0.239	0.071		0.424	0.213	
120	(1)	0.058	0.063	20	-0.001	0.003	21	0.485	0.496	24	0.999	0.987	48
	(2)	0.054	0.011		1.513	0.317		0.255	0.060		0.405	0.196	
							$k=2.0,\ \phi=0.9$						
30	(1)	0.414	0.327	33	-0.008	-0.004	27	0.875	0.888	35	0.990	0.970	53
	(2)	0.413	0.136		1.974	0.529		0.039	0.014		0.410	0.218	
60	(1)	0.289	0.190	28	0.002	0.001	22	0.884	0.894	28	0.997	0.981	49
	(2)	0.218	0.061		1.933	0.415		0.036	0.010		0.415	0.204	
120	(1)	0.230	0.123	21	-0.003	0.000	19	0.888	0.896	22	0.999	0.987	48
	(2)	0.172	0.036		1.690	0.325		0.052	0.011		0.463	0.221	
							$k=3.0,\ \phi=0.0$						
30	(1)	0.115	0.183	48	-0.005	-0.001	47	-0.032	-0.013	56	0.993	0.978	60
	(2)	0.237	0.114		2.178	1.020		0.318	0.177		0.344	0.205	
60	(1)	0.055	0.093	43	-0.004	-0.003	42	-0.014	-0.004	51	0.998	0.986	57
	(2)	0.122	0.052		2.040	0.854		0.345	0.175		0.337	0.192	
120	(1)	0.031	0.055	39	0.006	0.002	40	-0.009	-0.002	45	1.000	0.991	54
	(2)	0.060	0.024		2.106	0.841		0.330	0.148		0.335	0.180	
							$k=3.0,\ \phi=0.5$						
30	(1)	0.377	0.333	49	0.003	0.003	43	0.440	0.468	53	0.991	0.978	57
	(2)	0.471	0.229		2.599	1.127		0.258	0.137		0.345	0.198	
60	(1)	0.178	0.153	41	-0.003	0.000	41	0.472	0.488	43	0.995	0.985	56
	(2)	0.231	0.095		2.230	0.911		0.266	0.116		0.331	0.186	
120	(1)	0.093	0.083	40	0.001	0.000	38	0.486	0.495	42	0.997	0.989	53
	(2)	0.107	0.043		2.001	0.767		0.261	0.109		0.363	0.192	
							$k=3.0,\ \phi=0.9$						
30	(1)	0.459	0.411	53	-0.006	0.000	44	0.882	0.889	53	0.992	0.977	60
	(2)	0.611	0.323		2.652	1.176		0.025	0.013		0.321	0.192	
60	(1)	0.359	0.269	43	0.000	-0.002	41	0.887	0.893	42	0.995	0.983	55
	(2)	0.395	0.169		2.463	1.004		0.027	0.012		0.317	0.173	
120	(1)	0.262	0.189	42	0.010	0.007	40	0.891	0.895	41	1.001	0.993	54
	(2)	0.257	0.109		2.431	0.980		0.034	0.014		0.336	0.180	
							$k=5.0,\ \phi=0.0$						
30	(1)	0.214	0.269	67	0.002	0.003	65	-0.033	-0.020	72	0.994	0.986	70
	(2)	0.582	0.388		3.685	2.393		0.333	0.241		0.273	0.192	
60	(1)	0.085	0.122	62	-0.001	-0.002	64	-0.013	-0.006	66	0.996	0.991	66
	(2)	0.291	0.181		3.401	2.175		0.337	0.221		0.263	0.173	
120	(1)	0.063	0.082	61	-0.004	-0.001	61	-0.011	-0.006	64	1.001	0.996	65
	(2)	0.136	0.083		3.192	1.952		0.302	0.193		0.262	0.171	
							$k=5.0,\ \phi=0.5$						
30	(1)	0.550	0.515	69	-0.009	-0.008	64	0.446	0.463	72	0.995	0.985	69
	(2)	1.157	0.795		4.091	2.618		0.253	0.182		0.274	0.190	
60	(1)	0.320	0.299	66	0.001	0.002	63	0.468	0.478	67	0.998	0.991	64
	(2)	0.569	0.373		3.690	2.340		0.250	0.167		0.295	0.188	
120	(1)	0.144	0.138	61	-0.004	-0.001	59	0.485	0.491	65	1.000	0.995	63
	(2)	0.262	0.160		3.417	2.029		0.232	0.152		0.263	0.165	
							$k=5.0,\ \phi=0.9$						
30	(1)	0.503	0.510	68	0.008	0.005	63	0.888	0.891	69	0.991	0.984	67
	(2)	1.058	0.722		3.973	2.518		0.015	0.011		0.268	0.181	
60	(1)	0.384	0.368	65	0.005	-0.002	62	0.892	0.894	67	0.999	0.991	63
	(2)	0.727	0.470		3.888	2.417		0.017	0.012		0.274	0.174	
120	(1)	0.267	0.255	62	-0.016	-0.007	60	0.894	0.896	65	1.001	0.996	62
	(2)	0.509	0.314		3.828	2.302		0.023	0.015		0.269	0.168	

Table 1.(cont.ed.)

		$\widetilde{\mu}_i$	$\widehat{\mu}_i$	RE	$\widetilde{\tau}_i$	$\widehat{\tau}_i$	RE	$\widetilde{\phi}$	$\widehat{\phi}$	RE	$\widetilde{\sigma}$	$\widehat{\sigma}$	RE
$k = 10.0,\ \phi = 0.0$													
30	(1)	0.420	0.469	82	0.002	-0.001	81	-0.036	-0.029	87	0.994	0.991	81
	(2)	2.082	1.711		6.969	5.624		0.326	0.284		0.223	0.181	
60	(1)	0.178	0.215	81	0.001	0.000	80	-0.015	-0.011	82	0.997	0.994	78
	(2)	0.903	0.732		7.191	5.752		0.320	0.263		0.223	0.173	
120	(1)	0.105	0.138	76	0.005	0.005	80	-0.008	-0.006	80	0.998	0.995	77
	(2)	0.491	0.375		6.897	5.514		0.318	0.254		0.221	0.170	
$k = 10.0,\ \phi = 0.5$													
30	(1)	1.095	1.069	84	0.011	0.013	81	0.448	0.456	86	0.995	0.991	82
	(2)	3.973	3.322		8.117	6.609		0.223	0.191		0.222	0.181	
60	(1)	0.527	0.522	80	-0.012	-0.012	80	0.474	0.479	83	0.999	0.995	81
	(2)	2.103	1.684		7.936	6.369		0.234	0.191		0.228	0.184	
120	(1)	0.292	0.287	79	-0.010	-0.005	78	0.486	0.489	81	0.997	0.995	75
	(2)	0.948	0.753		7.838	6.137		0.227	0.185		0.223	0.167	
$k = 10.0,\ \phi = 0.9$													
30	(1)	0.570	0.636	86	0.008	0.009	82	0.893	0.894	86	0.995	0.990	82
	(2)	2.334	2.005		7.448	6.068		0.008	0.007		0.210	0.172	
60	(1)	0.441	0.476	81	-0.001	0.001	80	0.895	0.896	82	0.995	0.992	78
	(2)	1.573	1.275		7.433	5.924		0.009	0.008		0.218	0.170	
120	(1)	0.368	0.385	75	0.012	0.012	79	0.896	0.896	78	1.000	0.997	76
	(2)	1.223	0.920		7.554	6.001		0.014	0.011		0.213	0.161	

5. Power and robustness properties of the proposed test

For testing the null hypothesis $H_0 : \sum_{i=1}^{a} l_i\tau_i = \sum_{i=1}^{a} l_i\mu_i = 0$ ($\mu_i = \mu+\tau_i$); $\sum_{i=1}^{a} l_i = 0$, traditionally, where l_i $(1 \leq i \leq a)$ are constant coefficients of a linear contrast; we use the following test statistics based on the LS estimators given in 4.2

$$t = \frac{\sum_{i=1}^{a} l_i\widetilde{\mu}_i}{\sqrt{\sum_{i=1}^{a} l_i^2 \frac{\widetilde{\sigma}^2}{n}}}. \tag{5.1}$$

However, in this study, we propose the following test statistics based on MML estimators

$$t^* = \frac{\sum_{i=1}^{a} l_i\widehat{\mu}_i}{\sqrt{\sum_{i=1}^{a} l_i^2 \frac{\widehat{\sigma}^2}{m(k-1)}}}, \tag{5.2}$$

where the large values of t^* lead to the rejection of H_0. The null distribution of t^* is asymptotically normal N(0,1) due to the following lemmas:

5.1. Lemma. *For a given ϕ (σ known), the asymptotic distribution of $\widehat{\mu}_i(\phi,\sigma) = \widehat{\mu}_{i.} + (\Delta/m)\sigma$ which is the minimum variance bound estimator of $\mu_i = \mu + \tau_i$ $(1 \leq i \leq a)$ is normal with variance $V\{\widehat{\mu}_i(\phi,\sigma)\} \cong \sigma^2/m(k-1)$.*

Proof. Proof of the Lemma 5.1. The result follows from the fact that asymptotically $\partial lnL^*/\partial\mu_i$ is equivalent to $\partial lnL/\partial\mu_i$ [29] and assumes the form

$$\frac{\partial lnL^*}{\partial\mu_i} = \frac{m(k-1)}{\sigma^2}(\widehat{\mu}_i(\phi,\sigma) - \mu_i)$$

[10]. The normality follows from the fact that $E(\partial lnL^*/\partial\mu_i^r) = 0$ for all $r \geq 3$. □

5.2. Lemma. *For a given ϕ(μ known),the asymptotic distribution of $N\widehat{\sigma}^2(\phi,\mu)/\sigma^2$ is chi-square with $N = na$ degrees of freedom.*

Proof. Proof of the Lemma 5.2. Let

$$B_0 = (k-1)\sum_{i=1}^{a}\sum_{j=1}^{n}(y_{i,(j)} - \phi y_{i,(j-1)} - \mu_i)\Delta_j \quad \text{and}$$

$$C_0 = (k-1)\sum_{i=1}^{a}\sum_{j=1}^{n}\beta_j(y_{i,(j)} - \phi y_{i,(j-1)} - \mu_i)^2.$$

Since $B_0/\sqrt{nC_0} \cong 0$, α_j and β_j are bounded,

$$\begin{aligned}\frac{\partial lnL}{\partial\sigma} &\cong \frac{\partial lnL^*}{\partial\sigma}\\ &= -\frac{N}{\sigma^3}(\sigma - \frac{B_0+\sqrt{B_0^2+4NC_0}}{N})(\sigma - \frac{B_0-\sqrt{B_0^2+4NC_0}}{N})\\ &\cong \frac{N}{\sigma^3}(\frac{C_0}{N} - \sigma^2).\end{aligned}$$

The result then follows from the values of $E(\partial^r lnL^*/\partial\sigma^r)$ as in [20]. □

5.3. Lemma. *Since $\widehat{\sigma}$ converges to σ as n tends to infinity, the asymptotic distribution of $\sqrt{n/v_{11}}(\widehat{\mu}(\phi,\widehat{\sigma}) - \mu)/\widehat{\sigma}$ is N(0,1) where v_{11} is the first element in the asymptotic covariance matrix.*

Proof. Proof of the Lemma 5.3. This follows from the well-known Slutsky's theorem. See [20]. □

Thus, when we have a linear contrast of 'a' MML estimators and $\widehat{\sigma}^2$ is the pooled MML estimator of σ^2, the [2] conditions are satisfied and $\sum_{i=1}^{a} l_i\mu_i$ and $\widehat{\sigma}^2$ are asymptotically independently distributed resulting the asymptotic distribution of $\sqrt{m(k-1)}\sum_{i=1}^{a} l_i\widehat{\mu}_i/(\widehat{\sigma}\sqrt{\sum_{i=1}^{a} l_i^2})$ being N(0,1).

Some of the simulated values of the probabilities $P(t^* \geq z_{0.05} = 1.645|H_0)$ for different sample sizes are given in Table 2.

Table 2. Values of the type I error of the t^* test; $\alpha = 0.050$.

	$k = 2.0$	$k = 3.0$	$k = 5.0$	$k = 10.0$	$k = 15.0$
n			$\phi = 0.0$		
50	0.030	0.042	0.046	0.046	0.048
100	0.032	0.053	0.052	0.055	0.051
150	0.032	0.042	0.054	0.054	0.053
200	0.032	0.048	0.050	0.056	0.046
			$\phi = 0.4$		
50	0.034	0.047	0.054	0.058	0.048
100	0.030	0.045	0.047	0.053	0.040
150	0.030	0.041	0.053	0.051	0.056
200	0.036	0.046	0.054	0.052	0.040
			$\phi = 0.8$		
50	0.044	0.057	0.053	0.059	0.052
100	0.033	0.050	0.054	0.052	0.043
150	0.039	0.048	0.047	0.047	0.053
200	0.030	0.046	0.044	0.056	0.048

It can be seen that the normal distribution provides satisfactory approximations to the percentage points. To have an idea about the power of the two tests given in 5.1 and 5.2, the simulated values for $n = 100$ where $l_1 = 1$, $l_2 = -2$ and $l_3 = 1$ for different k and ϕ values are reported in Table 3. We carried out simulations for several other k, n and l_i values but did not report since they give the similar results.

The values of power given in Table 3 are obtained by adding a constant d to the observations in the first and the third treatments and subtracting $2d$ from the observations

Table 3. Values of the power of the t^* and t tests; $n = 100$.

	$k=2.0$		$k=3.0$		$k=5.0$		$k=10.0$		$k=15.0$	
d	t^*	t	t^*	t	t^*	t	t^*	t	t^*	t
					$\phi=0.0$					
0.000	0.035	0.040	0.035	0.044	0.034	0.041	0.044	0.066	0.056	0.065
0.013	0.120	0.091	0.100	0.103	0.094	0.088	0.100	0.093	0.088	0.086
0.025	0.304	0.175	0.232	0.166	0.170	0.140	0.152	0.153	0.135	0.137
0.038	0.589	0.257	0.382	0.234	0.332	0.277	0.259	0.239	0.257	0.244
0.050	0.774	0.331	0.583	0.347	0.469	0.356	0.378	0.330	0.357	0.321
0.063	0.920	0.467	0.735	0.471	0.600	0.487	0.540	0.468	0.473	0.437
0.075	0.966	0.591	0.866	0.589	0.759	0.630	0.658	0.589	0.606	0.557
0.088	0.993	0.704	0.929	0.716	0.860	0.698	0.772	0.703	0.723	0.683
0.100	0.999	0.795	0.979	0.787	0.920	0.780	0.865	0.785	0.839	0.789
0.113	1.000	0.865	0.994	0.888	0.963	0.874	0.906	0.864	0.890	0.856
					$\phi=0.4$					
0.000	0.026	0.052	0.035	0.048	0.045	0.059	0.040	0.056	0.048	0.057
0.013	0.082	0.090	0.110	0.113	0.084	0.077	0.067	0.067	0.085	0.078
0.025	0.225	0.130	0.189	0.139	0.160	0.130	0.166	0.164	0.159	0.154
0.038	0.383	0.188	0.322	0.220	0.276	0.223	0.258	0.237	0.264	0.245
0.050	0.623	0.267	0.487	0.322	0.404	0.306	0.372	0.332	0.363	0.338
0.063	0.816	0.368	0.613	0.373	0.549	0.427	0.522	0.441	0.498	0.458
0.075	0.888	0.474	0.758	0.483	0.685	0.520	0.605	0.541	0.609	0.584
0.088	0.960	0.543	0.869	0.594	0.802	0.611	0.752	0.681	0.708	0.676
0.100	0.981	0.607	0.935	0.701	0.874	0.739	0.833	0.768	0.816	0.771
0.113	0.996	0.726	0.967	0.769	0.927	0.781	0.906	0.857	0.894	0.866
					$\phi=0.8$					
0.000	0.035	0.054	0.042	0.058	0.040	0.053	0.044	0.043	0.051	0.049
0.013	0.124	0.086	0.125	0.094	0.118	0.103	0.133	0.105	0.115	0.102
0.025	0.357	0.163	0.308	0.210	0.265	0.214	0.246	0.223	0.239	0.223
0.038	0.620	0.271	0.508	0.294	0.457	0.348	0.438	0.380	0.410	0.392
0.050	0.848	0.405	0.705	0.461	0.638	0.499	0.603	0.541	0.621	0.564
0.063	0.951	0.543	0.874	0.580	0.792	0.643	0.777	0.690	0.793	0.739
0.075	0.988	0.685	0.956	0.757	0.920	0.788	0.904	0.830	0.902	0.849
0.088	1.000	0.768	0.990	0.849	0.963	0.879	0.951	0.919	0.955	0.931
0.110	1.000	0.858	0.998	0.927	0.993	0.944	0.988	0.964	0.980	0.969
0.113	1.000	0.913	0.998	0.947	0.997	0.973	0.998	0.987	0.994	0.988

in the second treatment. The results show that t^* test is much more powerful than the classical t test.

In practice, we may be in error when we assume that our data follow a particular distribution, since the shape parameters might be misspecified or the data might contain outliers, or be contaminated. When these situations arise, the distribution of the test statistic may differ from that expected. Therefore, the accurate estimates of the probability of type I and type II errors (i.e. power of the test) will not be obtained. When the underlying assumptions are violated, robust test statistics are preferred to the traditional test statistics. A test is called robust if its type I error is never substantially higher than a pre-assigned value for plausible alternatives to an assumed model (Criterion Robustness) and if its power is high (Inference Robustness). It is clear that robustness is very desirable property for the hypothesis testing procedures. Table 4 summarizes the results of simulations for $k = 3$, $\phi = 0.4$ and $n = 100$ when we assume that the true model is Gamma(3, σ). For this simulation study, the plausible alternatives used are as follows:

(1) Gamma(2,σ),
(2) Gamma(4,σ),
(3) Outlier model: $(n - r)$ observations come from Gamma(3,σ) but r observation (we do not know which one) comes from Gamma(3,2σ); $r = [0.5 + 0.1n]$,
(4) Mixture model: 0.90Gamma(3,σ) + 0.10Gamma(3,2σ),
(5) Contamination model: 0.90Gamma(3,σ) + 0.10Gamma(5,σ)

Table 4. Power of the t^* and t tests for alternatives to Gamma(3,σ); $k = 3$, $n = 100$ and $\phi = 0.4$.

	Model (1)		Model (2)		Model (3)		Moel (4)		Model (5)	
d	t^*	t	t^*	t	t^*	t	t^*	t	t^*	t
0.00	0.042	0.058	0.042	0.052	0.030	0.015	0.046	0.048	0.043	0.051
0.02	0.127	0.090	0.163	0.122	0.111	0.026	0.156	0.093	0.131	0.100
0.04	0.335	0.213	0.339	0.232	0.314	0.082	0.368	0.212	0.303	0.163
0.06	0.608	0.341	0.619	0.381	0.569	0.161	0.668	0.334	0.544	0.239
0.08	0.811	0.539	0.809	0.519	0.808	0.303	0.872	0.490	0.739	0.350
0.10	0.946	0.704	0.936	0.685	0.947	0.479	0.970	0.648	0.899	0.489
0.12	0.987	0.819	0.985	0.823	0.981	0.628	0.992	0.786	0.964	0.603

The values are obtained by adding a constant d to the observations in the first and the third treatments and subtracting $2d$ from the observations in the second treatment as in efficiency analysis. From Table 4, we see that the power of the t^* test is higher than the t test for all sample models given above. For sample models, except Model (3), in fact, the t^* test has a double advantage: not only has it much smaller type I error but also has higher power. Similar results are obtained for other ϕ values.

6. Determination of the shape parameter

It is known that when location, scale and shape parameters are to be estimated, maximum likelihood method is doubtful unless large samples ($n > 250$ or so) are available; see [6]. Thus, one should consider estimating location, scale or location and shape parameters when the sample size is small which is the case for experimental design. Therefore, in this study, it is assumed that the shape parameter k in 2.3 is known. Actually, an assumption of known shape parameter is found to be quite reasonable for many real-life problems; see for example, [9]. See also [13] for a better understanding of the importance of a given shape parameter.

However, in practice, shape parameter is also unknown. A plausible value for it can be identified by using Q-Q plots, goodness-of-fit tests, or by matching (approximately) the sample skewness and kurtosis with the corresponding values of the distribution. Also it can be determined by trying a series of values of this parameter as in [24]. The one that maximizes the likelihood function is the required estimate. Due to the intrinsic robustness of MMLE shown in section 5, this value will yield essentially the same estimates and standard errors for plausible alternatives.

7. Conclusion

In this study, we proposed a new test statistic for testing the assumed values of linear contrasts in one-way classification AR(1) model. We believe that the results of this study will be very useful for researchers and practitioners. Since all the procedures related with linear contrasts are based on the assumption of normality, homogeneity of variances and independence of error terms. There is a huge literature about nonnormality and heterogeneity of variances. However, there is no too much work when the independence assumption of error terms is not satisfied. Dependency is tried to be prevented at the design stage by randomization and there is a gap about how to deal with it, if it exists. This paper fills this gap not only by dealing with dependency but also with non-normality. The proposed test directly use the original data rather than the transformed data and is straightforward both algebraically and computationally.

Besides it has nice properties like efficiency and being robust to plausible deviations from the assumed model, i.e. not much affected from the outliers, contamination or the misspesification of the shape parameter. The robustness of the test is due to the half-umbrella ordering of the β_j coefficients, i.e. they decrease in the direction of the long tail(s). Thus, the extreme observations in the direction of the long tail(s) automatically receive small weights. That is instrumental to achieve robustness; see [8] and [20].

Acknowledgement

We are grateful to the reviewers for their comments to improve the manuscript.

References

[1] Akkaya, A.D. and Tiku, M.L. *Estimating parameters in autoregressive models in non-normal situations: Asymmetric innovations*, Commun. Statist.-Theory Meth. **30** (3), 517-536, 2001.

[2] Bartlett, M.S. *Approximate confidence intervals*, Biometrika **40**, 12-19, 1953.

[3] Bayrak, Ö.T. and Akkaya, A.D. *Estimating parameters of a multiple autoregressive model by the modified maximum likelihood method*, Journal of Computational and Applied Mathematics **233**, 1763-1772, 2010.

[4] Bhattacharyya, G.K. *The asymptotics of maximum likelihood and related estimators based on Type II censored data*, J. Amer. Statist. Assoc. **80**, 398-404, 1985.

[5] Bian, G. and Wong, W.K. *An alternative approach to estimate regression coefficients*, J. Appl. Statist. Sci. **6** (1), 21-44, 1997.

[6] Bowman, K.O. and Shenton, L.R. *Weibull distributions when the shape parameter is defined*, Computational Statistics and Data Analysis **36**, 299-310, 2001.

[7] Cryer, J.D. *Time Series Analysis* (PWS-KENT, 1986).

[8] Huber, P.J. *Robust Statistics* (John Wiley, 1981).

[9] Jun, C.H., Balamurali, S. and Lee, S.H. *Variables sampling plans for Weibull distributed lifetimes under sudden death testing*, IEEE Transactions on Reliability **55** (1), 53-58, 2006.

[10] Kendall, M.G. and Stuart, A. *The Advanced Theory of Statistics, Vol. 3* (McMillan, 1979).

[11] Li, W.K. and McLeod, A.I. *ARMA Modelling with non-Gaussian Innovations*, Journal of Time Series Analysis **9**, 155-168, 1988.

[12] Montgomery, D.C. *Introduction to Statistical Quality Control* (John Wiley, 1997).

[13] Nordman, D. and Meeker, W.Q. *Weibull prediction intervals for a future number of failures*, Technometrics **44**, 15-23, 2002.

[14] Puthenpura, S. and Sinha, N.K. *Modified maximum likelihood method for the robust estimation of system parameters from very noisy data*, Automatica **22**, 231-235, 1986.

[15] Sahai, H. and Ageel, M.I. *The Analysis of Variance: Fixed, Random and Mixed Models* (Birkhauser, 2000).

[16] Scheffe, H. *The Analysis of Variance* (John Wiley, 1959).

[17] Şenoğlu, B. *Robust 2^k factorial design with Weibull error distributions*, Journal of Applied Statistics **32** (10), 1051-1066, 2005.

[18] Şenoğlu, B. and Tiku, M.L. *Linear contrasts in experimental design with non-identical error distributions*, Biometrical J. **44**, 359-374, 2002.

[19] Tiku, M.L. *Estimating the mean and standard deviation from censored normal samples*, Biometrika **54**, 155-165, 1967.

[20] Tiku, M.L. and Akkaya, A.D. *Robust Estimation and Hypothesis Testing* (New Age International (P) Ltd., 2004).

[21] Tiku, M.L. and Wong, W.K. *Testing for a unit root in an AR(1) model using three and four moment approximations*, Commun. Stat.-Simula. **27**, 185-198, 1998.

[22] Tiku, M.L., Wong, W.K. and Bian, G. *Time series models with asymmetric innovations*, Commun. Statist.-Theory Meth. **28** (6), 1331-1360, 1999.

[23] Tiku, M.L., Wong, W.K. and Bian, G. *Estimating parameters in autoregressive models in non-normal situations: Symmetric innovations*, Commun. Statist.-Theory Meth. **28** (2), 315-341, 1999.

[24] Tiku, M.L., Wong, W.K., Vaughan, D.C. and Bian, G. *Time series models in non-normal situations: Symmetric innovations*, J. Time Series Analysis **21**, 571-596, 2000.

[25] Tiku, M.L. and Selçuk, A.S. *Robust time series: Some engineering applications*, in: IEEE Proceedings of Goa Conference, (Goa, India, 2000), 466-473.
[26] Türker, Ö. *Autoregressive Models: Statistical Inference and Applications*, PhD Thesis (METU: Turkey, 2002).
[27] Vaughan, D.C. *On the Tiku-Suresh method of estimation*, Commun. Statist.- Theory Meth. **21**, 391-404, 1992.
[28] Vaughan, D.C. *The generalized secant hyperbolic distribution and its properties*, Commun. Statist.-Theory Meth. **31**, 219-238, 2002.
[29] Vaughan, D.C. and Tiku, M.L. *Estimation and hypothesis testing for non-normal bivariate distribution and applications*, J. of Math. and Comp. Modelling **32**, 53-67, 2000.
[30] Vinod, H.D. and Shenton, L.R. *Exact moments for autoregressive and random walk models for a zero or stationary initial value*, Econometric Theory **12**, 481-499, 1996.
[31] Wong, W.K. and Bian, G. *Estimating parameters in autoregressive models with asymmetric innovations*, Statistics and Probability Letters **71**, 61-70, 2005.