Journal of Statistical Planning and Inference 142 (2012) 1027–1030

Contents lists available at SciVerse ScienceDirect

Journal of Statistical Planning and Inference

journal homepage: www.elsevier.com/locate/jspi

Short communication

# Comments on "Unbiased estimates for moments and cumulants in linear regression"

Iaroslav V. Blagouchine [a,*], Eric Moreau [b]

[a] Department of Mathematics, ISEN – Toulon & Telecommunication Department, ISITV of the University of Toulon, La Valette du Var (Toulon), France
[b] Telecommunication Department, ISITV of the University of Toulon, La Valette du Var (Toulon), France

ARTICLE INFO

*Article history:*
Received 26 August 2011
Received in revised form
16 October 2011
Accepted 20 October 2011
Available online 10 November 2011

*Keywords:*
Consistency
Cumulants
Linear regression
MA processes
Mixing
Moments
Non-i.i.d. processes
Semi-invariants
Unbiased estimates

ABSTRACT

In a recent number of the Journal of Statistical Planning and Inference, a short communication by Withers and Nadarajah (2011) on the unbiased estimation of the moments and cumulants for linear regression models was published. However, some of the results presented in this paper, as well as those closely related, were partially obtained in other works, which were not referenced or were referenced in other contexts. In addition, the consistency of the derived estimates was not studied. The aim of these short comments is to briefly present these results, as well as some omitted works, in order to give larger representation of the state of art in this field.



## 1. Introduction

In a recent paper by Withers and Nadarajah (2011), unbiased estimates (UEs) of the moments and cumulants for the stationary stochastic zero-mean processes in linear regression were considered. This paper contains some interesting results; nevertheless, we also noted that some aspects were not sufficiently developed. Namely, it concerns the bibliographical work, which was too sketchy, and the lack of studies related to certain basic performances of the derived estimators, such as consistency.

## 2. Bibliographical issues

In the introduction of the manuscript (Withers and Nadarajah, 2011), authors mentioned several previous works related to the subject. However, some important references were omitted or placed in other parts of the manuscript. For example, UEs for any order of cumulant for an i.i.d. process were given for the first time by Fisher (1929) (see also Kendall, 1945, Chapter 11); these estimators are known as $k$-statistics. Fisher's paper is mentioned in the work, but in other context

* Corresponding author. Tel.: +33 970 46 28 33; fax: +33 494 14 24 48.
*E-mail addresses:* iaroslav.blagouchine@univ-tln.fr (I.V. Blagouchine), moreau@univ-tln.fr (E. Moreau).
*URLS:* http://lseet.univ-tln.fr/~iaroslav/ (I.V. Blagouchine), http://moreau.univ-tln.fr/ (E. Moreau).

0378-3758/$ - see front matter 
doi:10.1016/j.jspi.2011.10.008

and in other part of the paper. Furthermore, the works of Kendall were not mentioned directly, albeit he greatly contributed to the theory of $k$-statistics. Authors only cited the book of Stuart and Ord (1987), which is a rewritten version of the original book of Kendall (1945) and in which the results of Kendall are discussed and presented in more modern form. One may additionally recall Kendall's (1940a–c) works, which could be mentioned in this context as well. Besides, a quite general paper on the estimation of semi-invariants[1] was published by Dressel (1940). Also, there are many interesting material regarding cumulant tensors and generalized $k$-statistics[2] published in the paper by Dwyer (1964), and especially, in the monograph of McCullagh (1987, Chapters 2–4).[3] Finally, in the book written by Ibragimov and Rozanov (1978, Chapter VII, Section 4), as well as in the paper by Comon (1992), estimation of linear regression coefficients in two different contexts is addressed. The latter problem is, in fact, the inverse one with respect to those considered by Blagouchine and Moreau (2010) and by Withers and Nadarajah (2011) respectively.

Now, the $k$-statistics are the UEs of cumulants elaborated in the hypothesis that we have absolutely no information about the underlying i.i.d. process and its moments. However, in practice, we often face zero-mean or/and known-variance processes, in which case unbiased $k$-statistics become less attractive because of greater value of the MSE. For such a case, Blagouchine and Moreau (2009) provided a simple method for the obtaining of the UEs for any order of cumulant, and gave explicit expressions of the UEs for the cumulants of order less than seven [section II-C, formulæ (11), (17)–(19)]. In addition, in this paper, authors particularly considered four different estimators of the fourth-order cumulant (three of which are unbiased) for an i.i.d. process. These estimators were worked out and studied in details in both batch and adaptive versions for the use with stationary and non-stationary stochastic processes respectively. The aforementioned paper (Blagouchine and Moreau, 2009) was cited by Withers and Nadarajah (2011), but in other context: first, it was referenced as giving UEs only for the fourth-order cumulant, while we also gave UEs for higher-order cumulants, and second, it was mentioned in the context of linear regression models, while in this publication only i.i.d. case was considered. On the contrary, in the later paper by Blagouchine and Moreau (2010), which was not referenced by Withers and Nadarajah (2011), authors worked out an unbiased estimator of the fourth-order cumulant for a stationary stochastic zero-mean finite MA process.[4]

## 3. Undiscussed consistency

Unfortunately, the consistency of the UEs derived by Withers and Nadarajah (2011) was not studied, nor even discussed. In contrast, in the aforementioned works, the consistency is studied, or at least, is discussed. For instance, in the fundamental paper of Fisher (1929), the consistency of the $k$-statistics for an i.i.d. process is straightforward because, first, the $k$-statistics is unbiased, and second, its variance, or second-order cumulant (denoted in works of Fisher and Kendall by $\kappa(r^2) \equiv \kappa_2[k_r] = \mathrm{var}[k_r]$, $r \in \mathbb{N}^*$, where $r$ is the order of the corresponding $k$-statistics, see e.g. Kendall (1945), p. 261), tends to zero for any $r$ when sample size $n \to \infty$.

In the paper (Blagouchine and Moreau, 2009), the consistency of the UEs for an i.i.d. process[5] is also straightforward since MSEs of all estimators approach zero as $n \to \infty$. In addition, in this paper, the consistency was also studied for the adaptive versions of the estimators, and convergences in mean and in mean-square were established. The question of consistency seems to be quite important for many other researchers in the field as well. Masry (2010) (another unreferenced paper) studied one of the unbiased estimators of the fourth-order cumulant from Blagouchine and Moreau (2009) for the $\rho$-mixing and strong-mixing processes. In particular, he showed that in these non-i.i.d. cases, studied estimator becomes biased, but is still consistent. Moreover, he also concluded, without proving, that similar method for the proof of consistency in $\rho$-mixing and strong-mixing contexts can be also applied to the UEs of the higher-order cumulants derived by Blagouchine and Moreau (2009).

As to the paper (Blagouchine and Moreau, 2010), the consistency of the proposed UE (see formulæ (8), (14)–(15)) was not proved theoretically: the unbiasedness was obtained analytically, while the MSE was studied only empirically. Notwithstanding, since finite MA processes are strong-mixing, the consistency of the UE from Blagouchine and Moreau (2010) follows directly from Masry's proof of consistency (given in Masry, 2010, Section II) for the UEs from Blagouchine and Moreau (2009). This may be shown as follows. We, first, recall that the $\rho$-mixing and strong-mixing processes, representing the broadest class of asymptotically independent processes, are defined in a following way: let $\{x_l\}$ be

[1] In literature, cumulants are formerly known as seminvariants (written also as semi-invariants) or half-invariants, see e.g. Thiele (1903) and Kendall (1945).

[2] Also known as multiple $k$-statistics, $l$-statistics or polykays, see e.g. Dwyer (1964).

[3] We also regret to not have mentioned some of these works in our earlier papers (Blagouchine and Moreau, 2009, 2010) (unfortunately, many engineering journals are quite strict in regard to the maximum length of the paper and to the scope of references).

[4] In statistical signal processing, where the notion of time is usually associated with the number of a random datum, linear regression models are more known as moving average (MA) processes with additional noise (which is the disturbance term in linear regression); see e.g. the equation in the abstract of Withers and Nadarajah (2011) and Blagouchine and Moreau (2010, Eq. (1)).

[5] Taking into account the maximum length of the paper and that our study was particularly devoted to the estimators of the fourth-order cumulant, the consistency was studied only for the estimators of the fourth-order cumulant.