Normal Approximation for Strong Demimartingales

Milto Hadjikyriakou*

Abstract

We consider a sequence of strong demimartingales. For these random objects, a central limit theorem is obtained by utilizing Zolotarev's ideal metric and the fact that a sequence of strong demimartingales is ordered via the convex order with the sequence of its independent duplicates. The CLT can also be applied to demimartingale sequences with constant mean. Newman (1984) conjectures a central limit theorem for demimartingales but this problem remains open. Although the result obtained in this paper does not provide a solution to Newman's conjecture, it is the first CLT for demimartingales available in the literature.

Key words and phrases: convex order, strong demimartingales, strong N-demimartingales, central limit theorem, Zolotarev's ideal metric.

AMS 2010 subject classification: 60F05, 62E17.

1 Introduction

Newman (1980) proved the following remarkable central limit theorem for associated random variables.

Theorem 1 *Let the sequence $\{X_n, n \geq 1\}$ be a strictly stationary associated sequence of random variables with $E(X_1^2) < \infty$ and $0 < \sigma^2 = \mathrm{Var}(X_1) + 2\sum_{j=2}^{\infty} \mathrm{Cov}(X_1, X_j) < \infty$. Then,*

$$\frac{S_n - E(S_n)}{\sqrt{n}} \xrightarrow{D} N(0, \sigma^2) \text{ as } n \to \infty$$

where $\xrightarrow{D}$ denotes convergence in distribution.

The result of Newman (1980) was the motivation for a number of central limit theorems for associated random variables (see for example Bulinski and Shaskin (2007), Prakasa Rao (2012), Oliveira (2012)).

Further to associated random variables, central limit theorems are provided for various notions of dependence such as martingales, mixing sequences, m-dependent random sequences, linearly positively/negatively quadrant dependent random variables (see

*School of Sciences, UCLAN Cyprus, 12-14 University Avenue, Pyla, 7080 Larnaka, Cyprus (email: miltwh@gmail.com).

for example Hall and Heyde (1980), Prakasa Rao (1975), Shang (2012), Boutsikas and Vaggelatou (2002)).

Newman and Wright (1982) introduced the concept of demimartingales in order to provide a much more general class than the associated random variables. The definition of demimartingales is given below.

Definition 2 *Let $\{S_n, n \geq 1\}$ be a collection of random variables defined on a probability space $(\Omega, \mathcal{A}, \mathcal{P})$. The sequence $\{S_n, n \geq 1\}$ is called a demimartingale if for every componentwise nondecreasing function f and for $j > i$*

$$E\left[(S_j - S_i)\, f(S_1, \ldots, S_i)\right] \geq 0 \tag{1}$$

If moreover (1) is valid for any nonnegative componentwise nondecreasing function f, then $\{S_n, n \geq 1\}$ is called a demisubmartingale.

Christofides and Hadjikyriakou (2015) introduced the concept of conditional strong demimartingales given a σ-field $\mathcal{F}$. The unconditional version of this definition is provided below.

Definition 3 *A sequence $\{S_n, n \in \mathbb{N}\}$ is said to be a strong demimartingale if for any two coordinatewise nondecreasing functions f and g and $j = 1, 2, \ldots$*

$$\operatorname{Cov}\left[g(S_{j+1} - S_j), f(S_1, \ldots, S_j)\right] \geq 0$$

whenever the covariance is defined.

Remark 4 *It can easily be proven that the partial sums of positively associated random variables form a sequence of strong demimartingales. Furthermore, if $\{S_n,\ n \in \mathbb{N}\}$ is a strong demimartingale sequence with $E(S_i) = C,\ \forall\ i = 1, 2, \ldots$ where C is a constant, the sequence $\{S_n,\ n \in \mathbb{N}\}$ is also a demimartingale.*

Concepts of dependence are closely related to stochastic orders. One of the most celebrated stochastic orders is the so-called convex order. A random variable X is said to be smaller than the random variable Y in the convex order (denoted by $X \preceq_{\mathrm{cx}} Y$) if $E\phi(X) \leq E\phi(Y)$ for all the convex functions ϕ such that the expectations exist (cf. Shaked and Shanthikumar (2007)).

Christofides and Hadjikyriakou (2015) proved a comparison theorem for conditionally strong demimartingales. The unconditional version of the theorem states that a sequence of strong demimartingales is always larger than the sequence of its independent duplicates in the convex order.

Theorem 5 *Let $\{S_n, n \in \mathbb{N}\}$ be a strong demimartingale and let $X_j = S_j - S_{j-1},\ j \geq 1$ with $S_0 \equiv 0$. Let X_j^* be independent random variables such that X_j and X_j^* have the same distribution and let $\hat{S}_n = \sum_{j=1}^{n} X_j^*$. Then,*

$$\hat{S}_n \preceq_{\mathrm{cx}} S_n.$$

Newman (1984) conjectures the following: Let $S_0 \equiv 0$ and the sequence $\{S_n, n \geq 1\}$ be an L^2-demimartingale whose difference sequence $\{X_n = S_n - S_{n-1},\ n \geq 1\}$ is strictly stationary and ergodic with

$$0 < \sigma^2 = \mathrm{Var}(X_1) + 2\sum_{j=2}^{\infty} \mathrm{Cov}(X_1, X_j) < \infty.$$

Then

$$n^{-1/2}(S_n - ES_n) \xrightarrow{D} \sigma N \text{ as } n \to \infty$$

where N is a standard normal random variable. The above conjecture has not been proven and the problem remains open. This paper aims to show that the result of Theorem 5 can be employed in order to obtain a central limit theorem for a class of strong demimartingales that is also valid for a class of demimartingale sequences. Although, the central limit theorem obtained in the next section does not provide a solution to Newman's conjecture to the best of my knowledge it is the first result in the literature dealing with the CLT for demimartingales.

2 Central limit theorem for strong demimartingales

The concepts of stochastic orders and probability metrics are closely related in the sense that if two random variables are somehow ordered and their expectations are close to one another, it is of interest to study how close their respective distributions are. In the case of random variables that are ordered with the convex order, a useful metric is the so called Zolotarev's ideal metric (Zolotarev (1983)) which is defined as

$$\zeta_s(X,Y) = \frac{1}{(s-1)!}\int_{-\infty}^{\infty} |E(X-t)_+^{s-1} - E(Y-t)_+^{s-1}|\ dt,\ \ s \in \mathbb{N} \setminus \{0\}$$

where $E|X|^{s-1} < \infty, E|Y|^{s-1} < \infty$ and $X_+ = \max\{0, X\}$.

Observe that if $X \preceq_{\mathrm{cx}} Y$ and $s = 2$ the above metric becomes of the form

$$\zeta_2(X,Y) = \int_{-\infty}^{\infty} (E(Y-t)_+ - E(X-t)_+)\ dt$$

where $E|X| < \infty, E|Y| < \infty$.

The main result of this paper is presented in this section and it is consider to be a Berry-Esseen type central limit theorem.

Theorem 6 *Let $\{S_n, n \in \mathbb{N}\}$ be a sequence of strong demimartingales and let $X_j = S_j - S_{j-1},\ j \geq 1$ be identically distributed with $S_0 \equiv 0$. Let $\mu = ES_1$, $\sigma^2 = E(S_1 - \mu)^2$ and assume that $E|S_1 - \mu|^3 < \infty$ and $\frac{1}{n}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} \mathrm{Cov}(X_i, X_j) \to 0$ as $n \to \infty$. Then*

$$\zeta_2\left(\frac{S_n - ES_n}{\sqrt{n}}, \mathcal{N}(0,\sigma^2)\right) \leq \frac{1}{n}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} \mathrm{Cov}(X_i, X_j) + \frac{c}{\sqrt{n}}(E|S_1 - \mu|^3 + 1) \qquad (2)$$

for a positive constant c.

Proof. Without loss of generality we assume that $ES_n = 0$ for $\forall n$. Let X_j^* and $\hat{S}_n$ be as stated in Theorem 5. Then $\hat{S}_n \preceq_{\text{cx}} S_n$ and by applying Theorem 4 of Boutsikas and Vaggelatou (2002) for $s = 2$ we have that

$$\begin{aligned}\zeta_2\left(\frac{S_n}{\sqrt{n}}, \frac{\widehat{S}_n}{\sqrt{n}}\right) &= \frac{1}{2n}\Big(\text{Var}(S_n) - \text{Var}(\widehat{S}_n)\Big) \\ &= \frac{1}{n}\sum_{1\le i<j\le n} \text{Cov}(X_i, X_j) \\ &= \frac{1}{n}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} \text{Cov}(X_i, X_j). \qquad (3)\end{aligned}$$

Since the sequence $\{X_n^*, n \in \mathbb{N}\}$ is i.i.d., by Theorem 4 of Senatov (1980) and for a constant $c_1 > 0$ we have

$$\zeta_2\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{X_i^*}{\sigma}, N\right) \le \frac{c_1}{\sqrt{n}}\left(\zeta_2\left(\frac{X_1}{\sigma}, N\right) + \max\left\{\zeta_1\left(\frac{X_1}{\sigma}, N\right), \zeta_3\left(\frac{X_1}{\sigma}, N\right)\right\}\right)$$

where $N \sim \mathcal{N}(0,1)$.
Observe that

$$E\left(\frac{X_1}{\sigma}\right) = 0 = E(N) \text{ and } E\left(\frac{X_1}{\sigma}\right)^2 = 1 = E(N^2).$$

Then by Proposition 2(iv) of Boutsikas and Vaggelatou (2002) for $s = 1, 2, 3$

$$\zeta_s\left(\frac{X_1}{\sigma}, N\right) \le \frac{1}{s!}\left(\frac{E|X_1|^s}{\sigma^s} + E|N|^s\right).$$

Note that

$$\zeta_1\left(\frac{X_1}{\sigma}, N\right) \le \frac{E|X_1|}{\sigma} + \sqrt{\frac{2}{\pi}} \le 1 + \sqrt{\frac{2}{\pi}},$$

$$\zeta_2\left(\frac{X_1}{\sigma}, N\right) \le 1$$

and

$$\zeta_3\left(\frac{X_1}{\sigma}, N\right) \le \frac{1}{6}\left(\frac{E|X_1|^3}{\sigma^3} + 2\sqrt{\frac{2}{\pi}}\right) = \frac{E|X_1|^3}{6\sigma^3} + \frac{1}{3}\sqrt{\frac{2}{\pi}}.$$

Hence

$$\begin{aligned}\zeta_2\left(\frac{\widehat{S}_n}{\sqrt{n}}, \mathcal{N}(0,\sigma^2)\right) &\le \frac{\sigma^2 c_1}{\sqrt{n}}\left(1 + \max\left\{1 + \sqrt{\frac{2}{\pi}}, \frac{E|X_1|^3}{6\sigma^3} + \frac{1}{3}\sqrt{\frac{2}{\pi}}\right\}\right) \\ &\le c\frac{E|X_1|^3 + 1}{\sqrt{n}} \qquad (4)\end{aligned}$$

where c is a positive constant. The desired result follows by using the triangular inequality and the relations (3) and (4). ■

It is worth mentioning that the result presented above provides rates of convergence in the CLT for strong demimartingales. If $n \to \infty$ the right hand side of inequality (2) tends to zero.

Next, we provide an example of a sequence of random variables for which all the assumptions of Theorem 6 are satisfied and therefore the CLT is applicable.

Example 7 *Let $X_1, X_2, \ldots, X_n$ be random variables from the normal distribution with mean equal to zero and variance equal to one. Furthermore assume that these random variables are positively correlated and therefore associated (Pitt (1982)).*

Suppose that the bivariate distribution of the vector (X_i, X_j) is given by the Farlie-Gumbel-Morgenstern system

$$F_{X_i,X_j}(x,y) = F(x)F(y)\left\{1 + \alpha_{ij}[1 - F(x)][1 - F(y)]\right\}. \tag{5}$$

where $F(x)$ is the common marginal cumulative distribution function of X_i's. It is known that for absolutely continuous random variables $|\alpha_{ij}| \leq 1$.

Schucany et al. (1978) proved that for the bivariate distribution described by (5) $Cov(X_i, X_j) = \frac{\alpha_{ij}}{\pi}$. Let α_{ij} be $\alpha_{ij} = \frac{1}{ij}$. Then

$$\begin{aligned} 0 &< \frac{1}{n}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} \mathrm{Cov}(X_i, X_j) \\ &\leq \frac{1}{n\pi}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} \frac{1}{ij} \\ &\leq \frac{1}{n\pi}(\ln n)^2 \\ &\to 0 \end{aligned}$$

as $n \to \infty$ which proves that all the assumptions of the previous theorem are satisfied and the CLT for the sequence S_n can be obtained.

Another celebrated metric that has been studied extensively is the uniform (or Kolmogorov) metric d_K which is defined as

$$d_K(X, Y) = \sup_{x \in \mathbb{R}} |F_X(x) - F_Y(x)|.$$

Rachev (1991) provides a relationship that links the Zolotarev's ideal metric ζ_2 to the Kolmogorov metric d_K. By using the result of Corollary 1 of Boutsikas and Vaggelatou (2002) and property 14.1.16 of Rachev (1991) we have that if $X \preceq_{\mathrm{cx}} Y$ or $Y \preceq_{\mathrm{cx}} X$ and Y has a bounded Lebesgue density f_Y, then

$$d_K(X, Y) \leq 3 \cdot 2^{-1/3} M_Y^{2/3} |EY^2 - EX^2|^{1/3} \tag{6}$$

where $M_Y = \sup_{x \in \mathbb{R}} f_Y(x)$.

By utilizing (6) we can obtain an inequality similar to (2) in terms of d_K. The proof follows by applying similar steps as in the proof of Theorem 6 and therefore is omitted for brevity.

Theorem 8 *Let S_n, $\widehat{S}_n$ and X_i be as stated in Theorem 5 and assume that $\widehat{S}_n$ has a bounded Lebesgue density f_Y. Let $\mu = ES_1$, $\sigma^2 = E(S_1-\mu)^2$ and assume that $E|S_1-\mu|^3 < \infty$ and $\frac{1}{n}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} \mathrm{Cov}(X_i, X_j) \to 0$ as $n \to \infty$. Then*

$$
\begin{aligned}
d_K\left(\frac{S_n - ES_n}{\sqrt{n}}, \mathcal{N}(0,\sigma^2)\right) \quad &\leq \quad \frac{3}{n} M_Y^{2/3} \left[\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} \mathrm{Cov}(X_i, X_j)\right]^{1/3} \\
&+ \quad \frac{c}{\sqrt{n}}\left(\frac{3}{\sqrt[3]{2\pi}} + \max\left\{1 + \sqrt{\frac{2}{\pi}}, \frac{E|X_1-\mu|^3}{6\sigma^3} + \frac{1}{3}\sqrt{\frac{2}{\pi}}\right\}\right)
\end{aligned}
$$

where c is a positive constant and $M_Y = \sup_{x\in\mathbb{R}} f_Y(x)$.

As it has already been mentioned a strong demimartingale $\{S_n, n \in \mathbb{N}\}$ with $E(S_n) = C$ for all n is also a demimartingale. Therefore, the result presented above provides a CLT for a sequence of demimartingales and as far as I am aware there are no results on central limit theorems for demimartingales in the literature.

The concept of strong N-demimartingales was introduced by Prakasa Rao (2004). The definition has a similar structure as the definition of strong demimartingales and it is given below.

Definition 9 *A sequence $\{S_n, n \in \mathbb{N}\}$ is said to be a strong N-demimartingale if for any two coordinatewise nondecreasing functions f and g and $j = 1, 2, \ldots$*

$$
\mathrm{Cov}\left[g(S_{j+1} - S_j), f(S_1, \ldots, S_j)\right] \leq 0.
$$

whenever the covariance is defined.

Hadjikyriakou (2013) provides a comparison theorem for the class of conditional strong N-demimartingales. By applying the unconditional version of the comparison theorem and similar same steps as in Theorem 6 with appropriate conditions the following result can also be obtained for a class of strong N-demimartingales.

Theorem 10 *Let $\{S_n, n \in \mathbb{N}\}$ be a sequence of strong N-demimartingales and let $X_j = S_j - S_{j-1},\ j \geq 1$ be identically distributed with $S_0 \equiv 0$. Let $\mu = ES_1$, $\sigma^2 = E(S_1-\mu)^2 < \infty$ and assume that $E|S_1-\mu|^3 < \infty$ and $\frac{1}{n}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} \mathrm{Cov}(X_i, X_j) \to 0$ as $n \to \infty$. Then*

$$
\zeta_2\left(\frac{S_n - ES_n}{\sqrt{n}}, \mathcal{N}(0,\sigma^2)\right) \leq \frac{c}{\sqrt{n}}(E|S_1-\mu|^3 + 1) - \frac{1}{n}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} \mathrm{Cov}(X_i, X_j) \qquad (7)
$$

for a positive constant c.

Again, since a sequence of strong N-demimartingales with constant means forms a sequence of N-demimartingales the latter inequality provides a CLT for a sequence of N-demimartingales.

Note that the results presented in Theorems 6 and 10 can also be used to provide central limit theorems for positively and negatively associated random variables respectively but of course better results for these classes of random variables are available in the literature. The contribution of this paper is the central limit theorem for strong demimartingales and strong N-demimartingales which are wider classes of random variables.

Acknowledgements

I am grateful to the anonymous referee for all the valuable comments which led to an improved version of the manuscript. I also would like to take the opportunity to thank Prof. T.C. Christofides for his constant support and attention to this work.

References

[1] Boutsikas, M.V., Vaggelatou, E. (2002): On the distance between convex-ordered random variables, with applications. *Advances in Applied Probability* **34**, 349-375. MR1909919

[2] Bulinski, A.V., Shaskin, A. (2007): Limit theorems for associated random fields of associated random variables. World Scientific, Singapore.

[3] Christofides, T.C., Hadjikyriakou, M. (2015): The Conditional Convex Order and a Comparison Inequality. *Stoch. Anal. Appl.* **33**, 259-270. MR3305469

[4] Esary, J. D., Proschan, F., Walkup, D. W. (1967): Association of random variables, with applications. *Ann. Math. Statist.* **38**, 1466-1474. MR0217826

[5] Hadjikyriakou, M. (2013). Comparison of conditional expectations of functions of strong N-demimartingales and functions of sums of conditionally independent random variables. *Statist. Probab. Lett.* **83**, 1282-1286. MR3041405

[6] Hall, P., and Heyde, C.C (1980). Martingale limit theory and its applications. Academic Press, Cambridge, MA.

[7] Newman, C.M. (1980): Normal Fluctuations and the FKG inequalities.*Comm. Math. Phys.* **95**, 75-80. MR0576267

[8] Newman, C. M.; Wright, A. L. (1982): Associated random variables and martingale inequalities. *Z. Wahrsch. Verw. Gebiete* **59** 361-371. MR0721632

[9] Newman, C.M. (1984): Asymptotic independence and limit theorems for positively and negatively dependent random variables. In: *Inequalities in Statistics and Probability*, Tong, Y.L. (ed.), 127-140 IMS, Hayward. MR0789244

[10] Oliveira, P.E. (2012): Asymptotics for Associated Random Variables. Springer, Berlin.

[11] Pitt, L. D. (1982). Positively correlated normal variables are associated. *The Annals of Probability*, 496-499.

[12] Prakasa Rao, B.L.S. (1975): Remark on the rate of convergence in the random central limit theorem for mixing sequences. *Probability Theory and Related Fields* **31**, 157-160. MR0370703

[13] Prakasa Rao, B.L.S. (2004): On some inequalities for N-demimartingales. *J. Indian Soc. Agricultural Statist.* **57**, 208-216. MR2070557

[14] Prakasa Rao, B.L.S. (2012): Associated Sequences, Demimartingales and Nonparametric Inference. Birkhäuser, Basel.

[15] Rachev, S.T. (1991): Probability metrics and the stability of stochastic models. John Wiley, New York.

[16] Schucany, W. R., Parr, W. C., and Boyer, J. E. (1978). Correlation structure in farlie-gumbel-morgenstern distributions. *Biometrika*, **65**, 650-653.

[17] Senatov, V.V. (1980): Uniform estimates of the rate of convergence in the multidimentional central limit theorem. *Theory of Probability and its Applications* **57**, 208-216. MR0595137

[18] Shaked, M. and Shanthikumar, J.G. (2007): Stochastic orders and their applications. Academic Press, New York.

[19] Shang, Y. (2012): A central limit theorem for randomly indexed m-dependent random variables. *Filomat* **26**, 713-717. MR3099010

[20] Zolotarev, V.M. (1983). Probability metrics. *Theory Prob. Appl.* **28**, 278-302.