Probability in the Engineering and Informational Sciences, **29**, 2015, 361–384.
doi:10.1017/S0269964815000042

A GENERAL CLASS OF CUSUM STATISTICS

CHRISTOPHER S. WITHERS
Industrial Research Limited
Lower Hutt
New Zealand
E-mail: kit.withers@gmail.com

SARALEES NADARAJAH
University of Manchester
Manchester M13 9PL
UK
E-mail: mbbsssn2@manchester.ac.uk

Much of the work relating to control charts is on monitoring observations for changes in mean. Here, for the first time, we develop a class of cusum-type statistics to monitor *any* aspect of a process (for example, mean, variance, skewness or kurtosis). This development is based on recently published papers by the authors.

1. INTRODUCTION

Monitoring of incoming observations for changes in mean via a cusum test statistic has received applications in many areas of the sciences. Set $X_0 = 0$ and let $X_1, X_2, \ldots$ be independent random variables in $\mathcal{R}^p$ from some cumulative distribution function $F(x)$ say, with mean μ, and finite moments. If $p = 1$ the rth cumulant of X_i denoted by $\kappa_r(X_i)$ is defined as the coefficient of $t^r/r!$ when $\log \mathbb{E}[\exp(tX_i)]$ is expanded as a power series. The joint cumulant of say $X_1, \ldots, X_n$ denoted by $\kappa(X_1, \ldots, X_n)$ is defined as

$$\kappa(X_1, \ldots, X_n) = \sum_{\pi} (|\pi| - 1)!(-1)^{|\pi|-1} \prod_{B \in \pi} \mathbb{E}\left(\prod_{i \in B} X_i\right),$$

where π runs through the list of all partitions of $\{1, \ldots, n\}$, B runs through the list of all blocks of the partition π, and $|\pi|$ is the number of parts in the partition. Set $\sigma^2 = \kappa_2$.

The observations $X_1, X_2, \ldots$ can be considered as a random process which may at some point go "out of control", changing their distribution. We define *the average process* of the observations as

$$M_n(t) = n^{-1} S_{[nt]} \tag{1.1}$$

for $S_0 = 0$, $S_i = \sum_{j=1}^{i} X_j$ and $0 \le t \le 1$, where $[x]$ is the integral part of x. A change in mean can be tracked by monitoring the average process via some functional of it, say

© Cambridge University Press 2015 0269-9648/15 $25.00

https://doi.org/10.1017/S0269964815000042 Published online by Cambridge University Press

$T(M_n)$, referred to as *a cusum statistic*. We denote the mean of the average process (1.1) by

$$m_n(t) = \mathbb{E}[M_n(t)] = \mu[nt]/n \to m(t) = \mu t$$

as $n \to \infty$ for $0 \le t \le 1$. Given a functional T, $\widehat{\theta} = T(M_n)$ can be thought of as an estimator of $\theta = T(m)$. In this way, we can if desired use $\widehat{\theta}$ to provide an estimator of μ.

We shall assume that the rth cumulant of $\widehat{\theta}$ has the standard expansion given by

$$\kappa_r\left(\widehat{\theta}\right) = \sum_{i=r-1}^{\infty} a_{r,i} n^{-i} \tag{1.2}$$

for $r = 1, 2, \ldots$. It follows from Withers [26] that for $a_{2,1} \neq 0$, then $Y_n = (n/a_{2,1})^{1/2}\left(\widehat{\theta} - \theta\right)$, where $\theta = a_{1,0}$, has Cornish–Fisher expansions of the form

$$P_n(x) = P(Y_n \le x) \approx \Phi(x) - \phi(x) \sum_{r=1}^{\infty} n^{-r/2} h_r(x), \tag{1.3}$$

$$p_n(x) = dP_n(x)/dx \approx \phi(x)\left\{1 + \sum_{r=1}^{\infty} n^{-r/2}\left(x h_r(x) - h_r'(x)\right)\right\}, \tag{1.4}$$

where $\Phi(x)$ is the cumulative distribution function of a unit normal random variable, $\phi(x)$ is the probability density function (pdf) of a unit normal random variable and $h_r(x)$ are certain polynomials in x and $A_{r,i} = a_{2,1}^{-r/2} a_{r,i}$, the standardized cumulant coefficients, listed in Withers [25], and $h_r'(x) = dh_r(x)/dx$. For example, $h_1(x) = A_{1,1} + A_{3,2}\left(x^2 - 1\right)/6$. In (1.3) and (1.4), $\Omega_n \approx \sum_{r=1}^{\infty} \omega_{n,r}$ means that there exist regularity conditions such that $|\Omega_n - \sum_{r=1}^{i} \omega_{n,r}| \to 0$ as $n \to \infty$ for every $i \ge 1$.

The usual nonparametric asymptotic methods for testing changes in mean give a first-order approximation in the one-sided case and a second-order approximation in the two-sided case, where by *Ith-order* we mean an absolute error of magnitude $n^{-I/2}$ for n the sample size. Withers and Nadarajah [27] showed how to improve on the order of these approximations for asymptotically normal linear statistics. Withers and Nadarajah [28] extended the work of Withers and Nadarajah [27] for non-linear statistics with non-zero asymptotic variances. This involved development of two calculi analogous to that of von Mises. Withers and Nadarajah [29] provided further extensions for non-linear statistics with zero asymptotic variances and for multivariate statistics.

However, there is relatively little work with respect to monitoring other measures of a process. Of the work we are aware of, we mention Chen and Gupta [8], Park, Lee, and Jeon [16], Andreou and Ghysels [3], Lee, Na, and Na [13], Wu and Tian [31], Zhang and Wu (2006), Wu, Zhang, and Wang [33], Davis, Lee, and Rodriguez-Yam [12], Aue et al. [4], Reynolds and Stoumbos [18], Shu, Yeung, and Jiang [20], Wu et al. [32], Castagliola and Maravelakis [7] and Withers and Nadarajah [30]. Most of these work are limited to monitoring of variance or monitoring of mean and variance. It appears that there is no work on monitoring of skewness, monitoring of kurtosis or monitoring of other measures. However, there is a large body of work on rank based process monitoring. It also appears that no one has considered monitoring of a *general* measure (a measure that does not take a specific form) of a process. Besides, we are not aware of any attempt to provide higher order approximations (that is, Ith-order approximations for *any* $I \ge 1$) for monitoring aspects

https://doi.org/10.1017/S0269964815000042 Published online by Cambridge University Press

other than mean. For quality control it is important, for example, that accurate tests are available for monitoring the variance as well as the mean.

There are many practical situations where monitoring of mean or variance is not sufficient. For example, Cushey, Bellin, and Rubin [11] say that "Estimates limited to only the mean and variance of the solute concentrations are often insufficient". Cascon, Keating, and Shadwick [6] argue that "point estimates of mean and variance of an assumed sampling distribution are insufficient summaries of the available information of future returns". In a microarray experiment for detection of significant genes, Wang et al. [22] find that "mean and variance contain insufficient information of the data when the sample size is small".

Monitoring aspects other than mean or variance is of interest in many practical situations. For example, monitoring of skewness arises in disturbances in the solar wind from interplanetary scintillations measurements (Rickett [19]); asset pricing (Neuberger [15]); turbulence in the interstellar medium from mergers, cool-core sloshing and jets (Vazza, Roediger, and Brüggen [21]). Monitoring of kurtosis arises in: lip detection and tracking (Caplier [5]); tool-condition monitoring from degradation signals (Chinnam, Rai, and Singh [9]); rolling element bearing faults detection (Al-Tubi and Al-Raheem [2]).

The aim of this paper is to extend the work of Withers and Nadarajah [27–29] for a class of cusum-type statistics to monitor *any* aspect of a process (for example, mean, variance, skewness or kurtosis). We also provide Ith-order approximations for monitoring of *any* aspect of a process for *any* $I \geq 1$. We believe that this is the first time cusum-type statistics have been developed to monitor *any* aspect of a process and to provide *any* order of approximation. The developments are described in Sections 2 and 3. Illustrations of these developments using simulated and real data are presented in Section 4. Some results concerning the Euler–McLaurin expansion needed are given in the Appendix. The results of this paper assume no particular forms for F.

In this paper, we shall not seek to compare our method with existing methods. It is difficult to see if this can be appropriate since (to the best of our knowledge) our method is the first to monitor *any* aspect of a process and to provide *any* order of approximation. In any case, we leave comparison with existing methods as future work. Other future work are to: (i) investigate optimality or asymptotic optimality of the statistics used in this paper; (ii) provide comparisons with other approaches for monitoring processes like minimizing the expected number of observations before an alarm is signalled.

Throughout this paper, one should interpret the right hand sides of equalities of the form $\Omega = \eta_{i,j,k,\ldots}$ as being implicitly summed over all possible values of $i, j, k, \ldots$. We shall also denote by $\omega^{(k)}(x)$ or $\omega_{.k}(x)$, the kth derivative of $\omega(x)$.

2. MONITORING OTHER ASPECTS OF THE DISTRIBUTION

In Withers and Nadarajah [27], we gave a method for monitoring the mean of F, μ, using the statistic $\int_0^1 W dM_n \xrightarrow{d} \mu \int_0^1 W$ as $n \to \infty$, where W is a given weight function. (We have switched from w to W to avoid confusion with w of Withers [23] below.) Suppose instead of μ we want to monitor some other aspect of F, say $T(F)$. The statistics we propose for this is

$$\widehat{\theta} = \int_{n^{-1}}^1 W(s) T\left(F_{[ns]}\right) ds = n^{-1} \sum_{j=1}^{n-1} W_{j,n} T\left(F_j\right) \xrightarrow{d} T(F) \int_0^1 W \qquad \textbf{(2.1)}$$

https://doi.org/10.1017/S0269964815000042 Published online by Cambridge University Press

as $n \to \infty$, where F_j is the empirical cumulative distribution function for a random sample of size j from F (if there is only one random sample available then F_j can be the empirical cumulative distribution function based on the first j observations) and

$$n^{-1}W_{j,n} = \int_{j/n}^{(j+1)/n} W.$$

We also consider its Studentized version $\widehat{\theta}_0$ below. ($\widehat{\theta}$ does not depend on X_n but this is not a problem as one can redefine the final sample by reducing n by 1.)

The choice for $W(t)$ depends on the scientific purpose for the control process. There is no rule of thumb for choosing $W(t)$. If more weight should be given to the later observations of the process then $W(t)$ should put more mass for t close to 1. If more weight should be given to the earlier observations of the process then $W(t)$ should put more mass for t close to 0. If more weight should be given to the middle observations of the process then $W(t)$ should put more mass for t close to 0.5. If there is no preference on the observations then $W(t)$ can be chosen to be a constant.

We shall see from Theorem 2.1 that $\widehat{\theta}$ satisfies the standard expansion (1.2). We assume $\widehat{\theta}$ is univariate. However, its multivariate extension (see Withers and Nadarajah [29]) will be obvious. To give the leading coefficients, set

$$\begin{aligned} m_j &= [ns_j], \\ m_0 &= \min(m_1, \ldots, m_r), \\ c_n(s_1, \ldots, s_r) &= n^{r-1}m_0/(m_1 \cdots m_r) \to c(s_1, \ldots, s_r) = \min(s_1, \ldots, s_r) / \prod_{j=1}^{r} s_j \end{aligned}$$

as $n \to \infty$, where s_j are any real numbers in $[0, 1]$. For example,

$$\begin{aligned} c_n(s_1, s_2) &= n/m_2 \text{ for } s_1 \le s_2, \\ c_n(s_1, s_2, s_3) &= n^2/(m_2 m_3) \text{ for } s_1 = \min(s_1, s_2, s_3). \end{aligned}$$

Set $W_n(s) = W(s)I(ns \ge 1)$, $W_{n,j} = W_n(s_j)$ and let $[\cdot]$ denote the notation of Withers [25]. For example,

$$\begin{aligned} [1^r] &= \int_{-\infty}^{\infty} T_F^r(x_1)\, dF(x_1), \\ [11] &= \int_{-\infty}^{\infty} T_F(x_1, x_1)\, dF(x_1), \\ [1, 12^3] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} T_F(x_1)\, T_F^3(x_1, x_2)\, dF(x_1)\, dF(x_2), \\ [1, 2, 3^2, 123] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} T_F(x_1)\, T_F(x_2)\, T_F^2(x_3) \\ &\quad \times T_F(x_1, x_2, x_3)\, dF(x_1)\, dF(x_2)\, dF(x_3), \end{aligned}$$

where $T_F(x_1)$ is the first von Mises derivative of $T(F)$ standardized so that $[1] = 0$, $T_F(x_1, x_2)$ is the second von Mises derivative of $T(F)$, and so on, and $T_F^r(x_1)$ is $T_F(x_1)$ raised to the power r, $T_F^3(x_1, x_2)$ is $T_F(x_1, x_2)$ raised to the power 3, and so on.

https://doi.org/10.1017/S0269964815000042 Published online by Cambridge University Press

THEOREM 2.1: *With the notation as above, $\widehat{\theta}$ satisfies (1.2) with*

$$a_{1,0} = \nu_{n,1,0} T(F),\ a_{2,1} = \nu_{n,2,1}\left[1^2\right],\quad a_{1,1} = \nu_{n,1,1}[11]/2, \tag{2.2}$$

$$a_{3,2} = \nu_{n,3,2}\left[1^3\right] + 3\widetilde{\nu}_{n,3,2}[1,12,2] \tag{2.3}$$

for

$$\begin{aligned}
\nu_{n,1,0} &= \int_0^1 W_{n,1} = \int_{n^{-1}}^1 W(s)ds,\\
\nu_{n,2,1} &= \int_0^1 \int_0^1 W_{n,1} W_{n,2} c_n(s_1,s_2)\, ds_1 ds_2\\
&= 2\int_0^1 W_{n,1} ds_1 (n/m_1) \int_0^{s_1} W_{n,2} ds_2,\\
\nu_{n,1,1} &= \int_0^1 W_{n,1} c_n(s_1,s_1)\, ds_1,\\
\nu_{n,3,2} &= \int_0^1 \int_0^1 \int_0^1 W_{n,1} W_{n,2} W_{n,3} c_n(s_1,s_2,s_3)\, ds_1 ds_2 ds_3\\
&= 3! \int_{\{n^{-1}\le s_1\le s_2\le s_3\le 1\}} W_{n,1} W_{n,2} W_{n,3} n^2 m_2^{-1} m_3^{-1} ds_1 ds_2 ds_3,\\
\widetilde{\nu}_{n,3,2} &= \int_0^1 \int_0^1 \int_0^1 W_{n,1} W_{n,2} W_{n,3} c_n(s_1,s_2)\, c_n(s_2,s_3)\, ds_1 ds_2 ds_3\\
&= 2n^2 \int_0^1 W_{n,1} ds_1 \int_0^{s_1} W_{n,3} ds_3 \Bigg\{ m_1^{-1} m_3^{-1} \int_0^{s_3} W_{n,2} ds_2\\
&\quad + m_1^{-1} \int_{s_3}^{s_1} W_{n,2} m_2^{-1} ds_2 + \int_{s_1}^1 W_{n,2} m_2^{-2} ds_2 \Bigg\}.
\end{aligned} \tag{2.4}$$

Theorem 2.1 gives expressions for $a_{1,0}$, $a_{2,1}$, $a_{1,1}$ and $a_{3,2}$. Expressions for $a_{2,2}$, $a_{4,3}$ are much longer and so are not given here. The proof of Theorem 2.1 is given later. An immediate consequence of Theorem 2.1 is the following.

THEOREM 2.2: *Set*

$$r_{i_1,\ldots,i_s} = \int_0^1 s_1^{-i_1} W(s_1)\, ds_1 \int_0^{s_2} \cdots \int_0^{s_{r-1}} s_r^{-i_r} W(s_r)\, ds_r. \tag{2.5}$$

Then, as $n \to \infty$, $\nu_{n,r,i} \to \nu_{r,i}$ and $\widetilde{\nu}_{n,3,2} \to \widetilde{\nu}_{3,2}$, where

$$\begin{aligned}
\nu_{1,0} &= \int_0^1 W = r_0,\\
\nu_{2,1} &= \int_0^1 \int_0^1 W(s_1) W(s_2) c(s_1,s_2)\, ds_1 ds_2 = 2r_{1,0},\\
\nu_{1,1} &= \int_0^1 W(s_1) s_1^{-1} ds_1 = r_1,
\end{aligned}$$

https://doi.org/10.1017/S0269964815000042 Published online by Cambridge University Press

$$\nu_{3,2} = \int_0^1 \int_0^1 \int_0^1 W(s_1) W(s_2) W(s_3) c(s_1, s_2, s_3) ds_1 ds_2 ds_3 = 3! \int_{0<s_1<s_2<s_3<1}$$
$$= 6 \int_0^1 W(s_3) s_3^{-1} ds_3 \int_0^{s_3} W(s_2) s_2^{-1} ds_2 \int_0^{s_2} W(s_1) ds_1 = 6 r_{1,1,0},$$
$$\widetilde{\nu}_{3,2} = \int_0^1 \int_0^1 \int_0^1 W(s_1) W(s_2) W(s_3) c(s_1, s_2) c(s_2, s_3) ds_1 ds_2 ds_3$$
$$= 2v_1 + 2v_2 = \nu_{3,2}/3 + 4r_{2,0,0}$$

for

$$v_1 = \int_{0<s_1<s_2<s_3<1} = \nu_{3,2}/6, \; v_2 = 2 \int_{0<s_1<s_3<s_2<1} = 2r_{2,0,0}. \tag{2.6}$$

Example 2.1: Take $W(s) = 1$. Then,

$$\nu_{n,1,0} = 1 - n^{-1}, \; \nu_{n,2,1} = 2\left(1 - n^{-1}\right) - n^{-1} \sum_{j=1}^{n-1} j^{-1}, \; \nu_{n,1,1} = \sum_{i=1}^{n-1} i^{-1},$$
$$\nu_{n,3,2} = 6\left(1 - n^{-1}\right) - 3n^{-1} \left(\sum_{i=1}^{n-1} i^{-1} \right)^2 /2.$$

So, $\nu_{1,0} = 1$, $\nu_{2,1} = 2$, $\nu_{1,1} = \infty$, $\nu_{3,2} = 6$. Also $\widetilde{\nu}_{3,2} = 5$. As $a_{1,1}$ is unbounded the expansions, (1.3) and (1.4), for $Y_n = (n/a_{2,1})^{1/2} \left(\widehat{\theta} - \theta\right)$ do not apply. But they do apply with $\theta = a_{1,0}$ replaced by $\theta = \mathbb{E}\left[\widehat{\theta}\right]$, giving $Y_n = Y_n'$ say. Since $\delta_n = n^{1/2}\left(\mathbb{E}\left[\widehat{\theta}\right] - a_{1,0}\right) = O\left(n^{-1/2} \log n\right) \to 0$, one can then obtain the distribution of $Y_n = n^{1/2}\left(\widehat{\theta} - a_{1,0}\right) = Y_n' + \delta_n$ by a Taylor series expansion about the cumulative distribution function of Y_n'.

Example 2.2: Take $W(s) = s$. Then,

$$\nu_{n,1,0} = \left(1 - n^{-2}\right)/2,$$
$$\nu_{n,2,1} = n^{-3} \sum_{i=1}^{n-1} \left(i^2 + 3i/2 - i^{-1}/4\right) = 1/3 + n^{-1}/4 - n^{-3}\left(s_{-1} + 7\right)/4,$$
$$\nu_{2,1} = 1/3,$$
$$\nu_{n,1,1} = 1 - n^{-1} + 2^{-1} n^{-1} \sum_{i=1}^{n-1} i^{-1} \to \nu_{1,1} = 1,$$
$$\nu_{3,2} = 1/4, \; \widetilde{\nu}_{3,2} = 19/48$$

for $s_{-1} = \sum_{i=1}^{n-1} i^{-1}$.

https://doi.org/10.1017/S0269964815000042 Published online by Cambridge University Press

Example 2.3: Take $T(F) = t(\mu)$, that is

$$\widehat{\theta} = n^{-1} \sum_{j=1}^{n-1} W_{j,n} t\left(\overline{X}_j\right),$$

where $\overline{X}_j$ is the mean of $X_1, \ldots, X_j$. Set $t_r = t^{(r)}(\mu)$. Then,

$$[1^r] = t_1^r \mu_r, \ [11] = t_2 \mu_2, \ [1,12,2] = t_1^2 t_2 \mu_2^2.$$

Now take $t(\mu) = \log \mu$, so that

$$a_{1,0} = \nu_{n,1,0} \log \mu, \ a_{2,1} = \nu_{n,2,1} \mu^{-2} \mu_2, \ a_{1,1} = -\nu_{n,1,1} \mu^{-2} \mu_2/2,$$
$$a_{3,2} = \nu_{n,3,2} \mu^{-3} \mu_3 - 3\widetilde{\nu}_{n,3,2} \mu^{-4} \mu_2^2,$$

where μ_r denotes the rth central moment. For the Central Limit Theorem one can replace $a_{1,0}$ and $a_{2,1}$ by their limits. This Central Limit Theorem agrees with that given for the case $W(s) = 1$ in Rempala and Wesolowski [17].

THEOREM 2.3: *The Studentized form of* $\widehat{\theta}$ *is*

$$\widehat{\theta}_0 = A\left(F_n\right)\left(\widehat{\theta} - a_{1,0}\right) \tag{2.7}$$

for $A(F) = \nu_{n,2,1}^{-1/2} V(F)^{-1/2}$ *and* $V(F) = \left[1^2\right]$. *Then, Eq. (2.1) in Withers and Nadarajah [28] holds with* $a_{r,i} = a_{r,i,0}$, *where in terms of* $a_{1,1}$, $a_{3,2}$ *of (2.2), (2.3),*

$$a_{1,0,0} = 0, \ a_{2,1,0} = 1,$$
$$a_{1,1,0} = \nu_{n,2,1}^{-1/2} \left[1^2\right]^{-1/2} a_{1,1} - \epsilon_n,$$
$$a_{3,2,0} = \nu_{n,2,1}^{-3/2} \left[1^2\right]^{-3/2} a_{3,2},$$
$$\epsilon_n = \nu_{n,2,1}^{-1/2} 2^{-1} \left[1^2\right]^{-3/2} \left(\left[1^3\right] + 2[1,12,2]\right) \nu_{n,1,1,0},$$
$$\nu_{n,1,1,0} = \int_{n^{-1}}^{1} W(s) ds.$$

So, $h_1(x) = A_{1,1} + A_{3,2}\left(x^2 - 1\right)/6$ *for* $\widehat{\theta}_0$ *is given in terms of* $h_1(x)$ *for* $\widehat{\theta}$, *where* $A_{1,1} = a_{2,1}^{-r/2} a_{1,1}$ *and* $A_{3,2} = a_{2,1}^{-r/2} a_{3,2}$.

The function $h_1(x)$ in Theorem 2.3 is needed for Edgeworth-type expansions for the distribution of $\widehat{\theta}$; see (1.3) and (1.4). We now provide the proofs of Theorems 2.1 and 2.3.

PROOF OF THEOREM 2.1: Set

$$\begin{aligned} \widehat{w}_{i,j} &= F_{M_j}\left(x_i\right), \ w_{i,j} = \mathbb{E}\left[\widehat{w}_{i,j}\right] = F\left(x_i\right), \\ S_j &= j/N, \ W_{N,j} = W_n\left(S_j\right), \\ M_j &= [nS_j], \ t\left(\widehat{w}_j\right) = T\left(F_{M_j}\right), \\ \widehat{\theta}_N &= t\left(\widehat{w}\right) = N^{-1} \sum_{j=1}^{N} W_{N,j} t\left(\widehat{w}_j\right), \end{aligned} \tag{2.8}$$

https://doi.org/10.1017/S0269964815000042 Published online by Cambridge University Press

$$c_{n,N}\left(S_{j_1},\ldots,S_{j_r}\right)=\min\left(M_{j_1},\ldots,M_{j_r}\right)/\left(M_{j_1}\cdots M_{j_r}\right),$$

$$t_{\cdot i_1,j_1,\ldots,i_r,j_r}=\partial^r t\left(\widehat{w}\right)/\partial\widehat{w}_{i_1,j_1}\cdots\partial\widehat{w}_{i_r,j_r}\Big|_{\widehat{w}=w},\ T_{i_1,\ldots,i_r}=T_F\left(x_{i_1},\ldots,x_{i_r}\right).$$

So, $\widehat{\theta}_N\xrightarrow{d}\widehat{\theta}$ as $N\to\infty$. For $k(\cdot)$ given by

$$k\left(x_1,\ldots,x_r\right)=\kappa\left(Y_1,\ldots,Y_r\right),$$

where $Y_i=I\left(X\le x_i\right)$ and $X\sim F(x)$ in $\mathcal{R}^p$, we can express the joint cumulant of $\left(\widehat{w}_{i_1,j_1},\ldots,\widehat{w}_{i_r,j_r}\right)$ as

$$\kappa\left(\widehat{w}_{i_1,j_1},\ldots,\widehat{w}_{i_r,j_r}\right)=n^{1-r}c_{n,N}\left(S_{j_1},\ldots,S_{j_r}\right)k\left(x_{i_1},\ldots,x_{i_r}\right)\tag{2.9}$$

since

$$\kappa\left(\sum_{i_1=1}^{\infty}a_{i_1,1}Z_{i_1},\ldots,\sum_{i_r=1}^{\infty}a_{i_r,r}Z_{i_r}\right)=\sum_{i=1}^{\infty}a_{i1}\cdots a_{ir}\kappa_r\left(Z_i\right)$$

for $\{Z_i\}$ independent. Now apply Withers [23] with $\widehat{w}_i$ replaced by $\widehat{w}_{i,j}$ and w_i replaced by $w_{i,j}=F\left(x_i\right)$. Then identify the partial derivatives of t with the von Mises derivatives, replacing single suffixes by double suffixes, $k_{r-1}^{i_1,\ldots,i_r}$ by $k_{r-1}^{i_1,j_1,\ldots,i_r,j_r}=c_{n,N}\left(S_{j_1},\ldots,S_{j_r}\right)k\left(dx_{i_1},\ldots,dx_{i_r}\right)$, other $k_j^{i_1,\ldots,i_r}$ by zero, and $\sum_{i,j,\ldots}$ by $\int\cdots\int dx_1dx_2\cdots$. So, (1.2) holds for $\widehat{\theta}_N$ with $a_{r,i}$ replaced by $a_{r,i,N}$, where $t_{i_1,\ldots,i_r}$ becomes

$$t_{\cdot i_1,j_1,\ldots,i_r,j_r}=I\left(j_1=\cdots=j_r\right)N^{-1}W_{N,j_1}T_{i_1,\ldots,i_r}.$$

So,

$$a_{1,0,N}=N^{-1}\sum_{j=1}^{N}W_{N,j}T(F)\to a_{1,0},$$

$$a_{2,1,N}=t_{\cdot i_1,j_1}k_1^{i_1,j_1,i_2,j_2}t_{\cdot i_2,j_2}=N^{-2}\sum_{j_1,j_2=1}^{N}W_{N,j_1}W_{N,j_2}c_{n,N}\left(S_{j_1},S_{j_2}\right)\left\langle 1^2\right\rangle,$$

$$a_{2,1,N}\to a_{2,1}=\nu_{n,2,1}\left\langle 1^2\right\rangle$$

as $N\to\infty$ for $\left\langle 1^2\right\rangle=\int_{-\infty}^{\infty}T_{i_1}T_{i_2}\ dk\left(x_{i_1},x_{i_2}\right)=\left[1^2\right]$. In the notation of Withers [23], $I_1\begin{pmatrix}2\\0\end{pmatrix}=0$ as $\mathbb{E}\left[\widehat{w}\right]=w$, so

$$2a_{1,1,N}=I_{0,1}\begin{pmatrix}2\\0\end{pmatrix}=t_{i,j}k_1^{i,j}=t_{\cdot i_1,j_1,i_2,j_2}k_1^{i_1,j_1,i_2,j_2}$$

$$=N^{-1}\sum_{j_1,j_2=1}^{N}I\left(j_1=j_2\right)W_{N,j_1}c_{n,N}\left(S_{j_1},S_{j_2}\right)\langle 11\rangle,$$

$$2a_{1,1,N}\to 2a_{1,1}=2\nu_{n,1,1}\langle 11\rangle$$

https://doi.org/10.1017/S0269964815000042 Published online by Cambridge University Press

as $N \to \infty$ for $\langle 11 \rangle = \int_{-\infty}^{\infty} T_{1,2} dk(x_1, x_2) = [11]$, and

$$I_3 \begin{pmatrix} 3 \\ 0 \end{pmatrix} = t_i t_j t_k k_2^{i,j,k} = t_{\cdot i_1, j_1} t_{\cdot i_2, j_2} t_{\cdot i_3, j_3} k_1^{i_1, j_1, i_2, j_2, i_3, j_3}$$

$$= N^{-3} \sum_{j_1, j_2, j_3 = 1}^{N} W_{N,j_1} W_{N,j_2} W_{N,j_3} c_{n,N}(S_{j_1}, S_{j_2}, S_{j_3})$$

$$\times \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} T_1 T_2 T_3 dk(x_1, x_2, x_3)$$

$$\to \nu_{n,3,2} \left[1^3\right]$$

as $N \to \infty$, and

$$I_{2,1} \begin{pmatrix} 2 & 2 \\ 0 & 0 \end{pmatrix} = t_i t_{j,k} t_\ell k_1^{i,j} k_1^{k,\ell}$$

$$= t_{\cdot i_1, j_1} t_{\cdot i_2, j_2, i_3, j_3} t_{\cdot i_4, j_4} k_1^{i_1, j_1, i_2, j_2} k_1^{i_3, j_3, i_4, j_4}$$

$$= N^{-3} \sum_{j_1, \ldots, j_4 = 1}^{N} W_{N,j_1} W_{N,j_2} I(j_2 = j_3) W_{N,j_4} c_{n,N}(S_{j_1}, S_{j_2}) c_{n,N}(S_{j_3}, S_{j_4}) \phi$$

$$= N^{-3} \sum_{j_1, j_2, j_3 = 1}^{N} W_{N,j_1} W_{N,j_2} W_{N,j_3} c_{n,N}(S_{j_1}, S_{j_2}) c_{n,N}(S_{j_2}, S_{j_3}) \phi$$

$$\to \widetilde{\nu}_{n,3,2} \phi$$

as $N \to \infty$, where $\phi = \int_{-\infty}^{\infty} T_1 T_{2,3} T_4 dk(x_1, x_2) dk(x_3, x_4) = [1, 12, 2]$. This completes the proof. ■

Proof of Theorem 2.3: Consider $\widehat{\theta}_0$ of (2.7). Identify $A(F_n)$ as $b(\widehat{w}_0)$, where $\widehat{w}_0 = \widehat{w}_N$. Replace $\widehat{\theta}_0$ by

$$\widehat{\theta}_{N0} = b(\widehat{w}_0) \left\{ N^{-1} \sum_{j=1}^{N} W_{N,j} t(\widehat{w}_j) - a_{1,0,N} \right\} = t_0(\widehat{w})$$

say for $a_{1,0,N} = N^{-1} \sum_{j=1}^{N} W_{N,j} t(w_j)$ and $W_{N,j}$, $\widehat{w}_j$, w_j as before for $i \geq 1$. Replace $\widehat{w}$ above by $\{\widehat{w}_j, 0 \leq j \leq N\}$. Note that (1.2) holds for $\widehat{\theta}_{N,0}$ with $a_{r,i}$ replaced by $a_{r,i,N,0}$, where

$$a_{2,1,N,0} = t_{0\cdot i} k_1^{i,j} t_{0\cdot j} = t_{0\cdot i_1, j_1} k_1^{i_1, j_1, i_2, j_2} t_{0\cdot i_2, j_2},$$

and so on. So,

$$a_{2,1,N,0} = t_{0\cdot 0}^2 k_1^{0,0} + 2 t_{0\cdot 0} \sum_{j=1}^{N} k_1^{0,j} t_{0\cdot j} + \sum_{i,j=1}^{N} t_{0\cdot i} k_1^{i,j} t_{0\cdot j},$$

$$t_{0\cdot 0} = 0.$$

https://doi.org/10.1017/S0269964815000042 Published online by Cambridge University Press

For $i \neq 0$, $t_{0\cdot i} = b(w_0)\, t_{\cdot i}$ for $t_{\cdot i}$ as before. So, $a_{2,1,N,0} = b(w_0)^2\, a_{2,1,N}$ for $b(w_0) = A(F)$, giving $a_{2,1,0} = A(F)^2 a_{2,1} = 1$. Similarly,

$$a_{1,0,0} = 0,\ a_{3,2,0} = A(F)^3 a_{3,2}.$$

For $a_{1,1,0}$, we need

$$I_{0,1}\begin{pmatrix}2\\0\end{pmatrix} = t_{0\cdot 0,0} k_1^{0,0} + 2\sum_{i=1}^{N} t_{0\cdot 0,i} k_1^{0,i} + \sum_{i,j=1}^{N} t_{0\cdot i,j} k_1^{i,j}.$$

The first term is zero as $t_{0\cdot 0,0} = 0$. The third term is $2A(F)a_{1,1}$. The second term is

$$2\int_{-\infty}^{\infty} A_0 T_1 dk\,(x_0, x_1)\,\nu_{n,1,1,0} = 2\int_{-\infty}^{\infty} A_1 T_1 dF_1 \nu_{n,1,1,0}$$

for $\nu_{n,1,1,0} = \int_{n^{-1}}^{1} w_1 c_n\,(1, s_1)\, ds_1$. Also $V(F) = \int_{-\infty}^{\infty} T_2^2 dF_2$, so

$$V_1 = T_1^2 - V(F) + 2\int_{-\infty}^{\infty} T_2\,(T_{1,2} - T_1)\, dF_2 = T_1^2 - V(F) + 2\int_{-\infty}^{\infty} T_2 T_{1,2} dF_2.$$

So, $\int_{-\infty}^{\infty} V_1 T_1 dF_1 = \left[1^3\right] + 2[1, 12, 2]$. So, the theorem follows. ■

If W is smooth enough at zero an alternative treatment is as follows. Let $a_{r,i}(T)$ denote $a_{r,i}$ for $T(F_n)$ given in Withers [25]: $a_{1,0}(T) = T(F)$, $a_{1,1}(T) = [11]/2$, and so on. So, $\mathbb{E}\left[T(F_j)\right] = \sum_{i=0}^{\infty} a_{1,i}(T) j^{-i}$. Set $u_{i,k}(t) = t^{-i} u_{\cdot k}(t)$, where $u(t) = \int_0^t W$. Then,

$$\mathbb{E}\left[\widehat{\theta}\right] = \sum_{j=0}^{\infty} f_j$$

for

$$f_j = \sum_{i=0}^{j} a_{1,i}(T) \sum_{m=0}^{j-i} \gamma_m\left(u_{i,k}\right)/k!\Big|_{k=j+1-m-i},$$
$$\gamma_m(g) = \alpha'_{1,m}(g) \text{ of (A.1)},\ m \neq 1,$$
$$\gamma_1(g) = -\left\{g(0) + g(1)\right\}/2.$$

This follows from (A.1) since

$$W_{j,n} n^{-1} = u\left((j+1)/n\right) - u(j/n) = \sum_{k=1}^{\infty} n^{-k} u_{\cdot k}(j/n)/k!.$$

https://doi.org/10.1017/S0269964815000042 Published online by Cambridge University Press

For example,

$$f_0 = T(F)\int_0^1 W,$$

$$f_1 = -W(0)T(F) + a_{1,1}(T)\int_0^1 t^{-1}W(t)\,dt,$$

$$f_2 = -3T(F)\{7W_{\cdot 1}(0) + 5W_{\cdot 1}(1)\} + a_{1,1}(T)\left\{\int_0^1 t^{-1}W_{\cdot 1}(t)\,dt/2 + \gamma_1(u_{1,2})\right\}$$
$$+ a_{1,2}(T)\int_0^1 t^{-2}W_{\cdot 1}(t)\,dt.$$

As j increases f_j is less likely to be finite. Higher cumulant coefficients can be written in terms of the joint cumulants of $\{T(F_j)\}$, obtainable by introducing N as above, then taking limits as $N \to \infty$.

3. A MORE GENERAL CLASS OF STATISTICS

Let us write the estimators $\widehat{\theta}$ of (2.1) as

$$\widehat{\theta}_1(W) = n^{-1}\int_0^1 W dM_n = n^{-1}\sum_{j=1}^{n} W(j/n)X_j,$$

$$\widehat{\theta}_2(W) = n^{-1}\sum_{j=1}^{n-1} W_{j,n}T(F_j).$$

Note that $\widehat{\theta}_2(W)$ is *not* a generalization from μ to $T(F)$ of $\widehat{\theta}_1(W)$, although both tend to $T(F)\int_0^1 W$ as $n \to \infty$. Nor is there a function W such that for $T(F) = \mu$, $\widehat{\theta}_2(W) \approx \widehat{\theta}_1(W)$ unless $W(1) = 0$. For if $T(F) = \mu$,

$$\widehat{\theta}_1(W) = n^{-1}\sum_{j=1}^{n-1} W(j/n)X_j,$$

$$\widehat{\theta}_2(W) = n^{-1}\sum_{j=1}^{n-1} b(j/n)X_j$$

for $b(t) = \int_t^1 s^{-1}W(s)ds$. Worse, if $W(s) > 0$ then $b(t) \to 0$ as $t \to 1$, whereas one would like to be able to choose the weight function to give more weight to the later observations.

Substituting $X_j = S_j - S_{j-1} = j\mu(F_j) - j\mu(F_{j-1})I(j \neq 1)$ into the expression for $\widehat{\theta}_1(W)$ gives

$$\widehat{\theta}_1(W) = n^{-1}\sum_{j=1}^{n} \lambda_{n,j}\mu(F_j)$$

for

$$\lambda_{n,j} = jW(j/n) - jW((j+1)/n)I(j \neq n). \tag{3.1}$$

https://doi.org/10.1017/S0269964815000042 Published online by Cambridge University Press

So, a natural extension of $\widehat{\theta}_1(W)$ from $\mu(F)$ to $T(F)$ is:

$$\widehat{\theta}'_\lambda = n^{-1}\sum_{j=1}^{n}\lambda_{n,j}T\left(F_j\right) \tag{3.2}$$

with the weights of (3.1). We first study the statistic

$$\widehat{\theta}_\lambda = n^{-1}\sum_{j=1}^{n-1}\lambda_{n,j}T\left(F_j\right), \tag{3.3}$$

where $\lambda_{n,j}$ are any weights (note (3.3) is more general than (2.1)). To obtain cumulant coefficients for $\widehat{\theta}_\lambda$, we can use the method and notation of the previous section with some slight changes:

THEOREM 3.1: *Set*

$$W_n(t) = \lambda_{n,[nt]}I(nt \geq 1),\ W_{n,j} = W_n\left(s_j\right),$$

$$\lambda_i = \lambda_{n,j_i},\ s_{i,k}(n-1) = \sum_{j_1=1}^{n-1} j_1^{-i}\lambda_1^k,$$

$$s_{i_r,k_r,\cdots,i_1,k_1}(n-1) = \sum_{j_r=1}^{n-1} j_r^{-i_r}\lambda_r^{k_r}\cdots\sum_{j_1=1}^{j_2-1} j_1^{-i_1}\lambda_1^{k_1},$$

$$r_{i_r,k_r,\cdots,i_1,k_1}(n) = \sum_{j_r=1}^{n} j_r^{-i_r}\lambda_r^{k_r}\cdots\sum_{j_1=1}^{j_2} j_1^{-i_1}\lambda_1^{k_1},$$

where, in the last line the summation is over $1 \leq j_1 \leq \cdots \leq n$, *and in the second to last line the summation is over* $0 < j_1 < \cdots < n$. *Then,*

$$\nu_{n,1,0} = n^{-1}\sum_{j=1}^{n-1}\lambda_{n,j} = n^{-1}s_{0,1}(n-1),$$

$$\nu_{n,2,1} = n^{-1}\sum_{j_1,j_2=1}^{n-1}\lambda_{n,j_1}\lambda_{n,j_2}\min\left(j_1,j_2\right)/\left(j_1j_2\right) = n^{-1}\left(2u_1+u_2\right) = n^{-1}\left(2u'_1-u_2\right)$$

for

$$u_1 = \sum_{j_1<j_2} = s_{1,1\cdot0,1}(n-1),\ u_2 = \sum_{j_1=j_2} = s_{1,2}(n-1),\ u'_1 = \sum_{j_1\leq j_2} = r_{1,1\cdot0,1}(n-1),$$

and

$$\nu_{n,1,1} = s_{1,1}(n-1),$$

$$\nu_{n,3,2} = n^{-1}\sum_{j_1,j_2,j_3=1}^{n-1}\lambda_{n,j_1}\lambda_{n,j_2}\lambda_{n,j_3}\min\left(j_1,j_2,j_3\right)/\left(j_1j_2j_3\right)$$

$$= n^{-1}\left(6u_1+3u_2+3u_3+u_4\right) = n^{-1}\left(6u'_1-3u'_2-3u'_3+7u_4\right)$$

https://doi.org/10.1017/S0269964815000042 Published online by Cambridge University Press

for

$$u_1 = \sum_{j_1<j_2<j_3} = s_{1,1\cdot1,1\cdot0,1}(n-1),\ u_2 = \sum_{j_1<j_2=j_3} = s_{2,2\cdot0,1}(n-1),$$
$$u_3 = \sum_{j_1=j_2<j_3} = s_{1,2\cdot1,2}(n-1),\ u_4 = \sum_{j_1=j_2=j_3} = s_{2,3}(n-1),$$
$$u'_1 = \sum_{j_1\le j_2\le j_3} = r_{1,1\cdot1,1\cdot0,1}(n-1) = u_1+u_2+u_3+u_4,$$
$$u'_2 = \sum_{j_1\le j_2=j_3} = r_{2,2\cdot0,1}(n-1) = u_2+u_4,$$
$$u'_3 = \sum_{j_1=j_2\le j_3} = r_{1,2\cdot1,2}(n-1) = u_3+u_4,$$

and

$$\widetilde{\nu}_{n,3,2} = n^{-1}\sum_{j_1,j_2,j_3=1}^{n-1}\lambda_{n,j_1}\lambda_{n,j_2}\lambda_{n,j_3}\min(j_1,j_2)\min(j_2,j_3)/\left(j_1j_2^2j_3\right)$$
$$= n^{-1}\left(u_1+2u_2+u_3\right) = n^{-1}\left(2u'_{1,1}-u'_{1,2}+4u_2-u'_{3,2}\right)$$

for

$$u_1 = \sum_{j_1\le j_2,j_3\le j_2} = \sum_{j_2=1}^{n-1} j_2^{-2}\lambda_{n,j_2}s_{0,1}(j_2)^2 = 2u'_{1,1}-u'_{1,2},$$
$$u'_{1,1} = \sum_{j_1\le j_3\le j_2} = r_{2,1\cdot0,1\cdot0,1}(n-1),\ u'_{1,2} = \sum_{j_1=j_3\le j_2} = s_{2,1\cdot0,2}(n-1),$$
$$u_2 = \sum_{j_1\le j_2<j_3} = \sum_{j_3=1}^{n-1} j_3^{-1}\lambda_{n,j_3}\sum_{j_2=1}^{j_3-1} j_2^{-1}\lambda_{n,j_2}s_{0,1}(j_2) = u'_{2,1}-u'_{2,2},$$
$$u'_{2,1} = r_{1,1\cdot1,1\cdot0,1}(n-1),\ u'_{2,2} = r_{2,2\cdot0,1}(n-1),$$
$$u_3 = \sum_{j_1<j_2,j_3} = \sum_{j_1=1}^{n-1} j_1^{-1}\lambda_{n,j_1}\sum_{j_3=1}^{n-1} j_3^{-1}\lambda_{n,j_3}s_{0,1}[\min(j_1,j_3)-1] = 2u_2-u'_{3,2},$$
$$u'_{3,2} = \sum_{j_2=j_3<j_1} = \sum_{j_2=j_3\le j_1} - \sum_{j_2=j_3=j_1} = r_{1,1\cdot1,2}(n-1)-r_{2,3}(n-1).$$

PROOF: Let $\widehat{\theta}_N$ and N be defined as in the proof of Theorem 2.1. As $N\to\infty$,

$$\widehat{\theta}_N \xrightarrow{d} \int_0^1 W_n(t)T\left(F_{[nt]}\right)\,dt = \widehat{\theta}_\lambda$$

of (3.3) and the cumulant coefficients of $\widehat{\theta}_\lambda$ are again given by (2.2)–(2.4). ■

https://doi.org/10.1017/S0269964815000042 Published online by Cambridge University Press

An alternative derivation of these results which does not need to introduce N then let $N \to \infty$, is to use

$$\widehat{w}_{i,j} = F_j(x_i),\ \widehat{w}_j = \{\widehat{w}_{i,j}\},\ \widehat{w} = \{\widehat{w}_j\},\ t_j(\widehat{w}_j) = T(F_j),$$

so

$$\widehat{\theta}_\lambda = n^{-1}\sum_{j=1}^{n-1}\lambda_{n,j}t_j(\widehat{w}_j) = t(w)$$

say. Then, (2.9) holds with $c_{n,N}$ replaced by c_n with $S_j = j/N$ and $c_n(S_{j_1},\ldots,S_{j_r}) = n^{r-1}M_0/(j_1\cdots j_r)$ for $M_0 = \min(j_1,\ldots,j_r)$. For third-order inference, we need $a_{2,2}$ and $a_{3,4}$ given by Theorem 3.2.

THEOREM 3.2: *Consider the estimator* $\widehat{\theta} = t(\widehat{w})$. *Set*

$$r_{i_1,k_1,\cdots,i_r,k_r} = \sum_{1\le j_1<n} j_1^{-i_1}\lambda_1^{k_1}\cdots\sum_{1\le j_r\le j_{r-1}} j_r^{-i_r}\lambda_r^{k_r},$$

and

$$s_{i_1,k_1,\cdots,i_r,k_r} = \sum_{1\le j_1<n} j_1^{-i_1}\lambda_1^{k_1}\cdots\sum_{1\le j_r< j_{r-1}} j_r^{-i_r}\lambda_r^{k_r}.$$

Then,

$$a_{2,2} = (r_{2,1\cdot 0,1} + r_{1,1\cdot 1,1} - r_{2,2})([1,11] + [1,122]) + (2r_{2,1\cdot 0,1} - r_{2,2})\left[12^2\right]/2 \qquad \textbf{(3.4)}$$

and

$$a_{3,4} = \sum_{i=1}^{4} c_k\nu_{n,3,4,k}J_k, \qquad \textbf{(3.5)}$$

where $c_k = 1, 12, 12, 4$ *for* $k = 1,\ldots,4$, *and*

$$\begin{aligned}
J_1 &= \left[1^4\right] - 3\left[1^2\right]^2,\\
\nu_{n,4,3,1} &= 24s_{1,1\cdot 1,1\cdot 1,1\cdot 0,1} + 12(s_{1,1\cdot 1,1\cdot 1,2} + s_{1,1\cdot 2,2\cdot 0,1} + s_{2,2\cdot 1,1\cdot 0,1})\\
&\quad + 4(s_{1,1\cdot 2,3} + s_{3,3\cdot 0,1}) + 6s_{2,2\cdot 1,2} + s_{3,4},\\
J_2 &= \left[1^2, 12, 2\right],\\
\nu_{n,4,3,2} &= n^{-1}\sum_{j_1,\ldots,j_4=1}^{n-1}\lambda_1\cdots\lambda_4\min(j_1,j_2,j_3)\min(j_3,j_4)/\left(j_1j_2j_3^2j_4\right),\\
J_3 &= [1, 12, 23, 3],\\
\nu_{n,4,3,3} &= n^{-1}\sum_{j_1,\ldots,j_4=1}^{n-1}\lambda_1\cdots\lambda_4\min(j_1,j_2)\min(j_2,j_3)\min(j_3,j_4)/\left(j_1j_2^2j_3^2j_4\right),\\
J_4 &= [1, 2, 3, 123],\\
\nu_{n,4,3,4} &= n^{-1}\sum_{j_1,\ldots,j_4=1}^{n-1}\lambda_1\cdots\lambda_4\min(j_1,j_2)\min(j_1,j_3)\min(j_1,j_4)/\left(j_1^3j_2j_3j_4\right).
\end{aligned}$$

https://doi.org/10.1017/S0269964815000042 Published online by Cambridge University Press

PROOF: Note that $a_{2,2}$ for $\widehat{\theta} = t(\widehat{w})$ is given by Withers [23] in the form $a_{2,2} = \sum_{k=1}^{5} c_k I_k$, where $c_k = 1, 1, 1/2, 1, 2$ for $k = 1, \ldots, 5$ and certain I functions: $I_k = \nu_{n,2,2,k} J_k$, where

$$
\begin{aligned}
J_1 &= 0, \ J_2 = [1, 11], \\
\nu_{n,2,2,2} &= n^{-2} \sum_{j_1, j_2 = 1}^{n-1} \lambda_1 \lambda_2 c_n (S_{j_1}, S_{j_2}, S_{j_2}) = \sum_{j_1, j_2 = 1}^{n-1} \lambda_1 \lambda_2 \min(j_1, j_2) / \left(j_1 j_2^2\right) \\
&= \sum_{j_1 < j_2} + \sum_{j_2 < j_1} + \sum_{j_1 = j_2} = s_{2,1\cdot 0,1} + s_{1,1\cdot 1,1} - s_{2,2} \\
&= \sum_{j_1 \le j_2} + \sum_{j_2 \le j_1} - \sum_{j_1 = j_2} = r_{2,1\cdot 0,1} + r_{1,1\cdot 1,1} - r_{2,2}, \\
2r_{1,1\cdot 1,1} &= r_{1,1}^2 - r_{2,2}, \\
J_3 &= \left[12^2\right], \\
\nu_{n,2,2,3} &= n^{-2} \sum_{j_1, j_2 = 1}^{n-1} \lambda_1 \lambda_2 c_n (S_{j_1}, S_{j_2})^2 = \sum_{j_1, j_2 = 1}^{n-1} \lambda_1 \lambda_2 \{\min(j_1, j_2) / (j_1 j_2)\}^2 \\
&= 2 \sum_{j_1 < j_2} + \sum_{j_1 = j_2} = 2 s_{2,1\cdot 0,1} + s_{2,2} \\
&= 2 \sum_{j_1 \le j_2} - \sum_{j_1 = j_2} = 2 r_{2,1\cdot 0,1} - r_{2,2}, \\
J_4 &= [1, 122], \\
\nu_{n,2,2,4} &= n^{-2} \sum_{j_1, j_2 = 1}^{n-1} \lambda_1 \lambda_2 c_n (S_{j_1}, S_{j_2}) c_n (S_{j_2}, S_{j_2}) \\
&= \sum_{j_1, j_2 = 1}^{n-1} \lambda_1 \lambda_2 \min(j_1, j_2) / \left(j_1 j_2^2\right) = \nu_{n,2,2,2}, \\
J_5 &= 0.
\end{aligned}
$$

The term $2I_5$ was omitted from Withers [24] but is given in its Appendix. So, (3.4) follows. Similarly, we have $a_{3,4} = \sum_{k=1}^{4} c_k \nu_{n,3,4,k} J_k$, where $c_k = 1, 12, 12, 4$ for $k = 1, \ldots, 4$, J_k as given in the statement of theorem, and

$$
\begin{aligned}
\nu_{n,4,3,1} &= n^{-4} \sum_{j_1, \ldots, j_4 = 1}^{n-1} \lambda_1 \cdots \lambda_4 c_n (S_{j_1}, \ldots, S_{j_4}) \\
&= n^{-1} \sum_{j_1, \ldots, j_4 = 1}^{n-1} \lambda_1 \cdots \lambda_4 \min(j_1, \ldots, j_4) / (j_1 \cdots j_4) \\
&= 4! \sum_{j_1 < \cdots < j_4} + \sum_{2j's \text{ the same}} + \sum_{3j's \text{ the same}} + \sum_{j_1 = \cdots = j_4} \\
&= 24 s_{1,1\cdot 1,1\cdot 1,1\cdot 0,1} + 12 (s_{1,1\cdot 1,1\cdot 1,2} + s_{1,1\cdot 2,2\cdot 0,1} + s_{2,2\cdot 1,1\cdot 0,1}) \\
&\quad + 4 (s_{1,1\cdot 2,3} + s_{3,3\cdot 0,1}) + 6 s_{2,2\cdot 1,2} + s_{3,4},
\end{aligned}
$$

https://doi.org/10.1017/S0269964815000042 Published online by Cambridge University Press

$$\nu_{n,4,3,2} = n^{-4} \sum_{j_1,\ldots,j_4=1}^{n-1} \lambda_1 \cdots \lambda_4 c_n\left(S_{j_1}, S_{j_2}, S_{j_3}\right) c_n\left(S_{j_3}, S_{j_4}\right)$$

$$= n^{-1} \sum_{j_1,\ldots,j_4=1}^{n-1} \lambda_1 \cdots \lambda_4 \min\left(j_1, j_2, j_3\right) \min\left(j_3, j_4\right) / \left(j_1 j_2 j_3^2 j_4\right),$$

$$\nu_{n,4,3,3} = n^{-4} \sum_{j_1,\ldots,j_4=1}^{n-1} \lambda_1 \cdots \lambda_4 c_n\left(S_{j_1}, S_{j_2}\right) c_n\left(S_{j_2}, S_{j_3}\right) c_n\left(S_{j_3}, S_{j_4}\right)$$

$$= n^{-1} \sum_{j_1,\ldots,j_4=1}^{n-1} \lambda_1 \cdots \lambda_4 \min\left(j_1, j_2\right) \min\left(j_2, j_3\right) \min\left(j_3, j_4\right) / \left(j_1 j_2^2 j_3^2 j_4\right),$$

$$\nu_{n,4,3,4} = n^{-4} \sum_{j_1,\ldots,j_4=1}^{n-1} \lambda_1 \cdots \lambda_4 c_n\left(S_{j_1}, S_{j_2}\right) c_n\left(S_{j_1}, S_{j_3}\right) c_n\left(S_{j_1}, S_{j_4}\right)$$

$$= n^{-1} \sum_{j_1,\ldots,j_4=1}^{n-1} \lambda_1 \cdots \lambda_4 \min\left(j_1, j_2\right) \min\left(j_1, j_3\right) \min\left(j_1, j_4\right) / \left(j_1^3 j_2 j_3 j_4\right).$$

So, (3.5) follows and the proof is complete. ■

These theorems extend to (3.2) simply by replacing the upper bound $n-1$ by n.

Example 3.1: Suppose $\lambda_{n,j} = W(j/n)$. As $n \to \infty$, $\nu_{n,r,i} \to \nu_{r,i}$ and $\widetilde{\nu}_{n,3,2} \to \widetilde{\nu}_{3,2}$ of (2.6). Set

$$e_{i,k} = n^{-i}\lambda_{n,n}^k = n^{-i}W(1)^k,\ W_{i,k}(t) = t^{-i}W(t)^k.$$

Then, using $|i| = i_1 + i_2 + \cdots$, $(f)_{1,n}$ of (A.8), t_n of (A.4) and (A.5) and α_k of (A.6) and (A.7), we have

$$s_{i,k}(n) = n^{1-i}\left(W_{i,k}\right)_{1,n},$$

$$r_{i_1,k_1\cdot i_2,k_2}(n-1) = r_{i_1,k_1\cdot i_2,k_2}(n) - n^{-i_1}W(1)^{k_1}s_{i_2,k_2}(n),$$

$$r_{i_1,k_1\cdot i_2,k_2}(n) = n^{2-i_1-i_2}t_n\left(W_{i_1,k_1}, W_{i_2,k_2}\right) = n^{2-|i|}\sum_{k=0}^{\infty} n^{-k}\alpha_k\left(W_{i_1,k_1}, W_{i_2,k_2}\right),$$

$$r_{i_1,k_1\cdot i_2,k_2\cdot i_3,k_3}(n) = n^{3-|i|}t_n\left(W_{i_1,k_1}, W_{i_2,k_2}, W_{i_3,k_3}\right) = n^{3-|i|}\sum_{k=0}^{\infty} n^{-k}\alpha_k\left(W_{i_1,k_1}, W_{i_2,k_2}, W_{i_3,k_3}\right).$$

So, in terms of $W_{1,1}(t) = t^{-1}W(t)$ and r of (2.5),

$$r_{1,1\cdot 1,1\cdot 0,1}(n) = nt_n\left(W_{1,1}, W_{1,1}, W\right) = \sum_{k=0}^{\infty} n^{1-k}\alpha_k,$$

$$r_{1,1\cdot 1,1\cdot 0,1}(n-1) = r_{1,1\cdot 1,1\cdot 0,1}(n) - n^{-1}W(1)r_{1,1\cdot 0,1}(n)$$

for

$$\alpha_0 = r_{1,1,0},$$

$$\alpha_1 = \int_0^1 W_{1,1}\int_0^t W_{1,1}\left\{W(t) - W(0)\right\}dt/2 + \int_0^1 W_{1,1}^2\int_0^t W/2 + W(1)r_{1,0}.$$

https://doi.org/10.1017/S0269964815000042 Published online by Cambridge University Press

So, for smooth W,

$$\begin{aligned}
\nu_{n,1,0} &= r_0 - n^{-1}\{W(0)+W(1)\}/2 + O\left(n^{-2}\right),\\
\nu_{n,2,1} &= \sum_{i=0}^{\infty} \nu_{2,1\cdot i} n^{-i} \text{ for } \nu_{2,1\cdot 0} = 2g_0\left(W_{1,1}, W\right) = 2r_{1,0}\\
&= n^{-1}\{2n - 3\Psi(n+1) - 3\gamma\},\\
\nu_{2,1\cdot 1} &= \int_0^1 W(t)\{W(t) - W(0)\}t^{-1}dt - 2W(1)\int_0^1 W - \int_0^1 W(t)^2 t^{-1}dt,\\
\nu_{n,3,2} &= 1/3\Bigg\{18n - 9\Psi(n+1) - 9\gamma - 9\gamma\Psi(n+1) - 9\gamma^2\\
&\quad - 18\sum_{j=1}^{n}\frac{\Psi(j)}{j} + 6\Psi'(n+1) - \pi^2 + 9\sum_{k=1}^{n}\frac{\Psi(k)}{k}\Bigg\}n^{-1},\\
\widetilde{\nu}_{n,3,2} &= 1/3\Bigg\{3n + 3\Psi(n+1) + 3\gamma - 3\sum_{j=1}^{n}\frac{\Psi(j)}{j} + 6\Psi'(n+1)\\
&\quad - \pi^2 + 3\sum_{j=1}^{n}\frac{\Psi(j)}{j}\Bigg\}n^{-1},
\end{aligned}$$

and so on, where γ is Euler's constant, $\Psi(x) = d\log\Gamma(x)/dx$ is the digamma function, and $\Psi'(x) = d\Psi(x)/dx$ is the trigamma function. In general $\nu_{n,r,i}$ does not have a power series in n^{-1}.

From the cumulant coefficients for $\widehat{\theta}_\lambda$ above we obtain the coefficients for $\widehat{\theta}_\lambda(\alpha) = \alpha T(F_n) + \widehat{\theta}_\lambda$ for a given constant α by adding $N\alpha$ to $W_{N,N}$ of (2.8). So, the cumulant coefficients of $\widehat{\theta}_\lambda(\alpha)$ are given by (2.2), (2.3) with $\nu_{n,r,i}$ and $\widetilde{\nu}_{n,3,2}$ replaced by $\nu'_{n,r,i}$ and $\widetilde{\nu}'_{n,3,2}$, respectively, where

$$\begin{aligned}
\nu'_{n,1,0} &= \nu_{n,1,0} + \alpha,\\
\nu'_{n,2,1} &= \nu_{n,2,1} + 2\alpha\nu_{n,1,0} + \alpha^2,\\
\nu'_{n,1,1} &= \nu_{n,1,1} + \alpha,\\
\nu'_{n,3,2} &= \nu_{n,3,2} + 3\alpha\nu_{n,2,1} + 3\alpha^2\nu_{n,1,0},\\
\widetilde{\nu}'_{n,3,2} &= \widetilde{\nu}_{n,3,2} + \alpha\left(2\nu_{n,2,1} + \nu_{n,1,0}^2\right) + 3\alpha^2\nu_{n,1,0} + \alpha^3.
\end{aligned}$$

To apply Withers and Nadarajah ([27], Section 2) to get second-order inference from these expressions, one needs to have $\left\{\nu_{n,r,i}, \widetilde{\nu}_{n,3,2}, \nu'_{n,r,i}, \widetilde{\nu}'_{n,3,2}\right\}$ bounded as $n \to \infty$.

4. ILLUSTRATION

Here, we perform a simulation study and analyze a real data set to illustrate the results in Section 2, in particular Theorem 2.1.

We take $X_1, X_2, \ldots$ to be independent normal random variables with unit mean and unit variance. We take $t = 1$. We take $T(M_n)$ as the skewness or the coefficient of variation.

https://doi.org/10.1017/S0269964815000042 Published online by Cambridge University Press

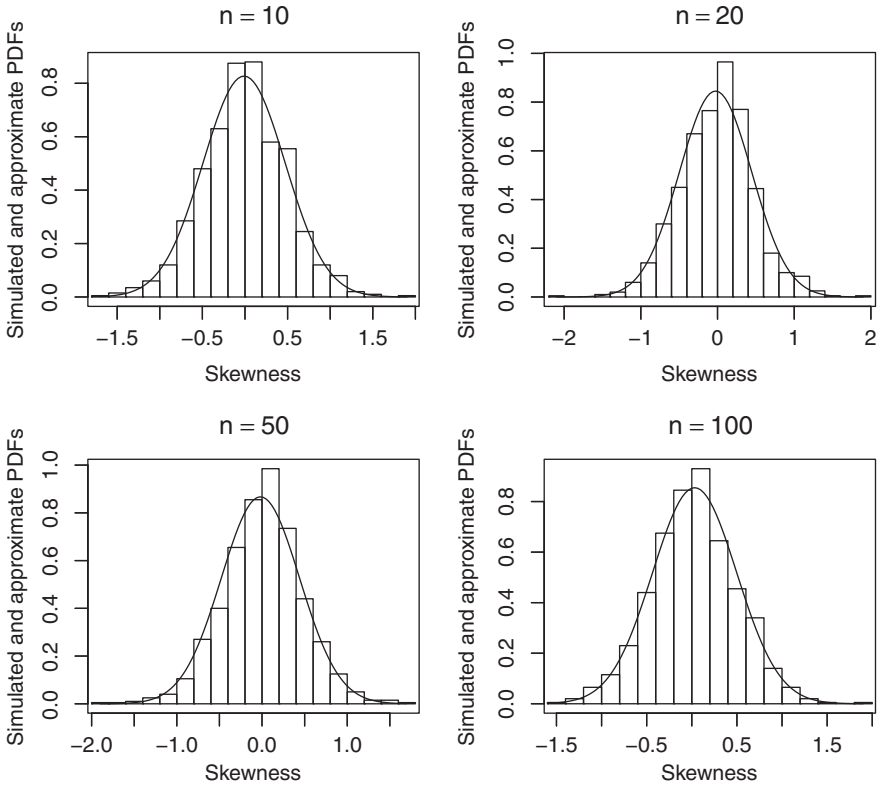


FIGURE 1. Simulated and approximate pdfs of the skewness.

We simulated the distribution of $T(M_n)$ as follows: (i) simulate twenty random samples $X_1, X_2, \ldots, X_n$ from the normal distribution with unit mean and unit variance; (ii) compute the corresponding M_n; (iii) use the twenty M_n values in (ii) to compute $T(M_n)$; (iv) repeat steps (i) to (iii) one thousand times. The simulated pdfs of $T(M_n)$ are shown in Figures 1 and 2 for $n = 10, 20, 50, 100$. Also shown in the figures are the third-order approximate pdfs of $T(M_n)$ obtained using Theorem 2.1. The approximation appears reasonable. As expected, the approximation appears to improve with increasing n.

The real data set that we use is due to Montgomery [14 pp. 183–184]. In a textile finishing plant, dyed cloth is inspected for the occurrence of defects per fifty square meters. The data are the number of non-conformities per $50\,\mathrm{m}^2$: 14, 12, 20, 11, 7, 10, 21, 16, 19, 23.

We simulated the distribution of $T(M_n)$ as follows: (i) resample 20 random samples $X_1, X_2, \ldots, X_4$ from the observed data; (ii) compute the corresponding M_n; (iii) use the 20 M_n values in (ii) to compute $T(M_n)$; (iv) repeat steps (i)–(iii) 100 times, making sure that the resampled random samples in (i) are independent each time. The resampled pdfs of $T(M_n)$ are shown in Figures 3 and 4. Also shown in the figures are the third-order approximate pdfs of $T(M_n)$ obtained using Theorem 2.1 with the population quantites replaced by sample estimates. The approximation appears reasonable again.

In the real data application, we have supposed that the approximations in Section 2 hold when the population quantities are replaced by sample estimates. This may not be the case. A future work is to investigate the effect of sampling errors on the approximations. It is pleasing however that the approximations in Figures 3 and 4 appear reasonable.

https://doi.org/10.1017/S0269964815000042 Published online by Cambridge University Press

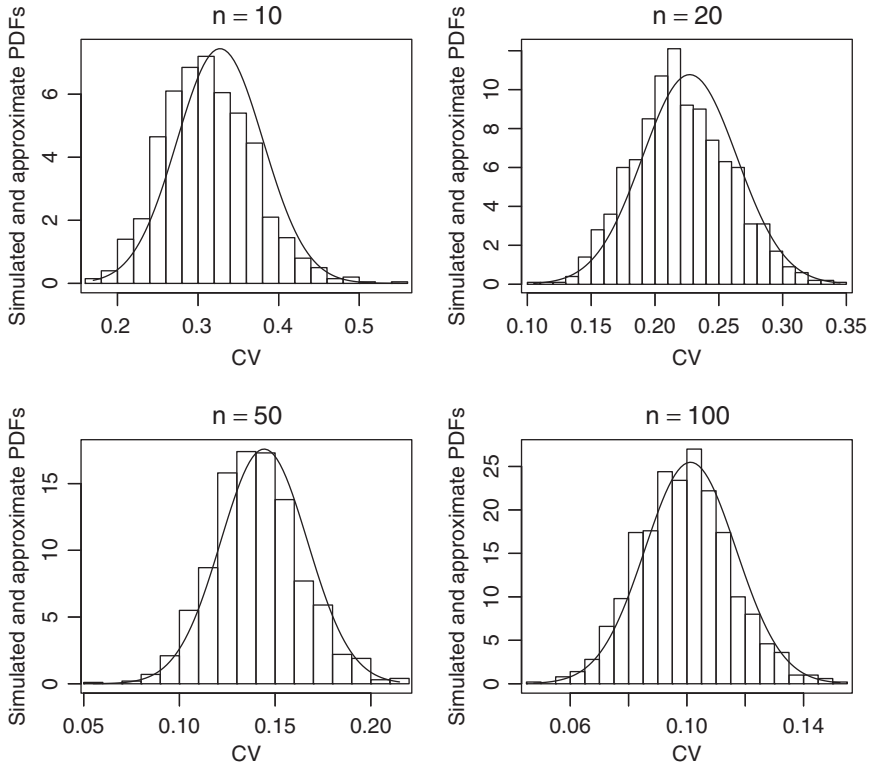


FIGURE 2. Simulated and approximate pdfs of the coefficient of variation.

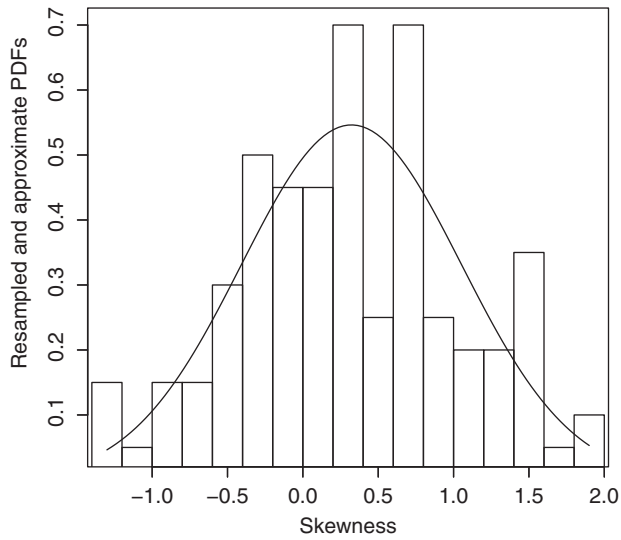


FIGURE 3. Resampled and approximate pdfs of the skewness for the real data set.

https://doi.org/10.1017/S0269964815000042 Published online by Cambridge University Press

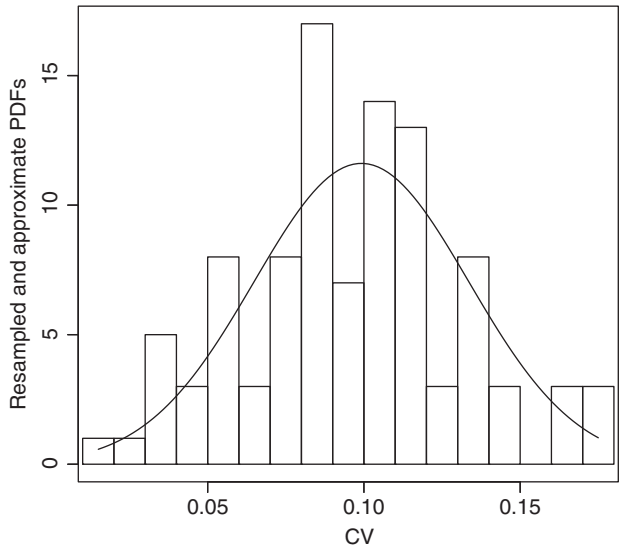


FIGURE 4. Resampled and approximate pdfs of the coefficient of variation for the real data set.

Acknowledgments

The authors would like to thank the Editor, the Associate Editor and the four referees for careful reading and for comments which greatly improved the paper.

References

1. Abramowitz, M. & Stegun, I.A. (1964). *Handbook of mathematical functions*. United States Department of Commerce, National Bureau of Standards, Applied Mathematics Series, vol. 55.
2. Al-Tubi, M.A.S. & Al-Raheem, K.F. (2010). Rolling element bearing faults detection, A time-domain analysis. *Caledonian Journal of Engineering* 6: 44–48.
3. Andreou, E. & Ghysels, E. (2003). Tests for breaks in the conditional co-movements of asset returns. *Statistica Sinica* 13: 1045–1073.
4. Aue, A., Hörmann, S., Horváth, L. & Reimherr, M. (2009). Break detection in the covariance structure of multivariate time series models. *Annals of Statistics* 37: 4046–4087.
5. Caplier, A. (2001). Lip detection and tracking. In: *Proceedings of the 11th International Conference on Image Analysis and Processing*, volume 4393, pp. 22–29.
6. Cascon, A., Keating, C. & Shadwick, W.F. (2003). The Omega function. Working paper.
7. Castagliola, P. & Maravelakis, P.E. (2011). A CUSUM control chart for monitoring the variance when parameters are estimated. *Journal of Statistical Planning and Inference* 141: 1463–1478.
8. Chen, J. & Gupta, A.K. (1997). Testing and locating variance changepoints with application to stock prices. *Journal of the American Statistical Association* 92: 739–747.
9. Chinnam, R.B., Rai, B. & Singh, N. (2004). Tool-condition monitoring from degradation signals using Mahalanobis–Taguchi system analysis. In: *Robust Engineering* pp. 343–351.
10. Csörgo, M. & Révész, P. (1981). *Strong approximations in probability and statistics*. New York: Academic Press.
11. Cushey, M.A., Bellin, A. & Rubin, Y. (1995). Three-dimensional modelling of flow and reactive transport in heterogeneous porous media. In: *Groundwater Quality—Remediation and Protection*, IAHS Publication no. 225, pp. 257–261.
12. Davis, R.A., Lee, T.C.M. & Rodriguez-Yam, G. A. (2008). Break detection for a class of nonlinear time series models. *Journal of Time Series Analysis* 29: 834–867.
13. Lee, S., Na, O. & Na, S. (2003). On the cusum of squares test for variance change in nonstationary and nonparametric time series models. *Annals of the Institute of Statistical Mathematics* 55: 467–485.

https://doi.org/10.1017/S0269964815000042 Published online by Cambridge University Press

14. Montgomery, D.C. (1991). *Introduction to statistical quality control*. 2nd ed. New York: John Wiley and Sons.
15. Neuberger, A. (2011). *Realized skewness*. Coventry, UK: Warwick Business School, Warwick University.
16. Park, S., Lee, S. & Jeon, J. (2000). The cusum of squares test for variance changes in infinite order autoregressive models. *Journal of the Korean Statistical Society* 29: 351–360.
17. Rempala, G. & Wesolowski, J. (2002). Asymptotics for products of sums and U-statistics. *Electronic Communications in Probability* 7: 47–54.
18. Reynolds, M.R. & Stoumbos, Z.G. (2010). Robust CUSUM charts for monitoring the process mean and variance. *Quality and Reliability Engineering International* 26: 453–473.
19. Rickett, B.J. (1975). Disturbances in the solar wind from IPS measurements in August 1972. *Solar Physics* 43: 237–247.
20. Shu, L., Yeung, H.-F. & Jiang, W. (2010). An adaptive CUSUM procedure for signaling process variance changes of unknown sizes. *Journal of Quality Technology* 42: 69–85.
21. Vazza, F., Roediger, E. & Brüggen, M. (2012). Turbulence in the ICM from mergers, cool-core sloshing and jets: results from a new multi-scale filtering approach.
22. Wang, J., Jia, M., Zhu, L., Yuan, Z., Li, P., Chang, C., Luo, J., Liu, M. & Shi, T. (2010). Systematical detection of significant genes in microarray data by incorporating gene interaction relationship in biological systems. *PLoS ONE* 5: e13721.
23. Withers, C.S. (1982a). The distribution and quantiles of a function of parameter estimates. *Annals of the Institute of Statistical Mathematics*, A 34: 55–68.
24. Withers, C.S. (1982b). Second order inference for asymptotically normal random variables. *Sankhyā*, B 44: 1–9.
25. Withers, C.S. (1983). Expansions for the distribution and quantiles of a regular functional of the empirical distribution with applications to nonparametric confidence intervals. *Annals of Statistics* 11: 577–587.
26. Withers, C.S. (1984). Asymptotic expansions for distributions and quantiles with power series cumulants. *Journal of the Royal Statistical Society* B 46: 389–396.
27. Withers, C.S. & Nadarajah, S. (2009). Accurate tests and intervals based on linear cusum statistics. *Statistics and Probability Letters* 79: 689–697.
28. Withers, C.S. & Nadarajah, S. (2009). Accurate tests and intervals based on non-linear cusum statistics. *Statistics and Probability Letters* 79: 2242–2250.
29. Withers, C.S. & Nadarajah, S. (2012). Accurate tests and intervals based on multivariate cusum statistics. *Sequential Analysis* 31: 78–87.
30. Withers, C.S. & Nadarajah, S. (2013). Weighted cusums for increased power near the end points. *Methodology and Computing in Applied Probability* 15: 379–405.
31. Wu, Z. & Tian, Y. (2005). Weighted-loss-function CUSUM chart for monitoring mean and variance of a production process. *International Journal of Production Research* 43: 3027–3044.
32. Wu, Z., Yang, M., Khoo, M.B.C. & Yu, F.-J. (2010). Optimization designs and performance comparison of two CUSUM schemes for monitoring process shifts in mean and variance. *European Journal of Operational Research* 205: 136–150.
33. Wu, Z., Zhang, S. & Wang, P. (2007). A CUSUM scheme with variable sample sizes and sampling intervals for monitoring the process mean and variance. *Quality and Reliability Engineering International* 23: 157–170.
34. Zhang, S. & Wu, Z. (2006). Monitoring the process mean and variance using a weighted loss function CUSUM scheme with variable sampling intervals. *IIE Transactions* 38: 377–387.

APPENDIX

Here, we give the Euler–McLaurin expansion, Eq. (23.1.30), page 806 of Abramowitz and Stegun [1], and related results. The main results are Propositions A.1–A.4.

PROPOSITION A.1: *Suppose that* $g : [0, 1] \to \mathcal{R}$ *has finite derivatives of order* $2K$*, where* $K \geq 1$ *is a fixed integer. Then, for* $r = 1$*, we have the expansions*

$$n^{-1} \sum_{i=0}^{n-1} g(i/n) = \sum_{k=0}^{2K-1} \alpha'_k(g) n^{-k} + r_{n,2K}, \tag{A.1}$$

https://doi.org/10.1017/S0269964815000042 Published online by Cambridge University Press

$$n^{-1}\sum_{i=1}^{n} g(i/n) = \sum_{k=0}^{2K-1} \alpha_k(g) n^{-k} + r_{n,2K}, \tag{A.2}$$
$$n^{-1}\sum_{i=1}^{n-1} g(i/n) = \sum_{k=0}^{2K-1} \alpha''_k(g) n^{-k} + r_{n,2K},$$

where

$$\alpha'_0(g) = \int_0^1 g(t)\,dt,$$
$$\alpha'_k(g) = \left\{g^{(k-1)}(1) - g^{(k-1)}(0)\right\} B_k/k! \text{ for } k = 1, 2, 3, \ldots,$$
$$\alpha_1(g) = \{g(1) - g(0)\}/2 = -\alpha'_1(g),$$
$$\alpha''_1(g) = -\{g(0) + g(1)\}/2,$$
$$\alpha_k(g) = \alpha''_k(g) = \alpha'_k(g) \text{ for } k \neq 1,$$
$$r_{n,k} = n^{-k}\left(B_k/k!\right) n^{-1}\sum_{j=0}^{n-1} g^{(k)}\left(j/n + \theta/n\right),\ 0 < \theta < 1,$$

and B_k is the kth Bernoulli number.

Tables of Bernoulli numbers are given by the left column in page 809 of Abramowitz and Stegun [1]. Note that (A.2) implies that

$$\int_0^1 g(t) d\ell_n(t) = \sum_{k=0}^{2K-2} \alpha_{k+1}(g) n^{-k} + n r_{n,2K}$$

for $\ell_n(t) = [nt] - nt$.

Example A.1: *For f smooth,*

$$\sum_{k=1}^{n} \{f(k/n) - f((k-1)/n)\} = \left[f^2 - f f_{\cdot 1}\right]_0^1 + O\left(n^{-1}\right).$$

In contrast for f a Wiener process, the left-hand side tends to one almost surely as $n \to \infty$ via $n = \log_2 \mathcal{N}$ and $\mathcal{N}$ an integer: see page 83 of Csörgo and Révész [10].

Proposition A.2 is a more general result than Proposition A.1.

Proposition A.2: *For f smooth and $m > n \geq 0$,*

$$h\sum_{k=0}^{m-1} f(a+kh) = \sum_{k=0}^{2K-1} \alpha'_k(a, a+mh, f) h^k + r_{2K}(h, 0, m, f),$$
$$h\sum_{k=n}^{m-1} f(a+kh) = \sum_{k=0}^{2K-1} \alpha'_k(a+nh, a+mh, f) h^k + r_{2K}(h, n, m, f),$$
$$h\sum_{k=1}^{m} f(a+kh) = \sum_{k=0}^{2K-1} \alpha_k(a, a+mh, f) h^k + r_{2K}(h, 0, m, f),$$
$$h\sum_{k=n+1}^{m} f(a+kh) = \sum_{k=0}^{2K-1} \alpha_k(a+nh, a+mh, f) h^k + r_{2K}(h, n, m, f), \tag{A.3}$$

https://doi.org/10.1017/S0269964815000042 Published online by Cambridge University Press

where

$$\alpha_0'(a,b,f) = \int_a^b f(t)\,dt,$$
$$\alpha_k'(a,b,f) = \left\{f^{(k-1)}(b) - f^{(k-1)}(a)\right\} B_k/k! \text{ for } k \neq 0,$$
$$\alpha_k(a,b,f) = \alpha_k'(a,b,f) \text{ for } k \neq 1,$$
$$\alpha_1(a,b,f) = \{f(b) - f(a)\}/2 = -\alpha_1'(a,b,f),$$
$$r_k(h,n,m,f) = h^k\,(B_k/k!)\,h\sum_{j=n}^{m-1} f^{(k)}\,(a + jh + \theta h)\,,\ 0 < \theta < 1.$$

Note that if $f^{(k)}$ is bounded in $[0,1]$ then $r_k(h,n,m,f) = O\left(h^{k+1}(m-n)\right)$ and $r_{n,k} = O\left(n^{-k}\right)$.

For simplicity, take $K = \infty$. Then the following simpler result can be obtained.

PROPOSITION A.3: *Set*

$$t_n\,(w_1,w_2) = n^{-2}\sum_{j_1=1}^{n} w_2\,(j_1/n)\sum_{j_2=1}^{j_1} w_1\,(j_2/n)\,, \tag{A.4}$$

$$t_n\,(w_1,w_2,w_3) = n^{-3}\sum_{j_1=1}^{n} w_3\,(j_1/n)\sum_{j_2=1}^{j_1} w_2\,(j_2/n)\sum_{j_3=1}^{j_2} w_1\,(j_1/n)\,. \tag{A.5}$$

Then,

$$t_n\,(w_1,w_2) = \sum_{k=0}^{\infty} n^{-k}\alpha_k\,(w_1,w_2) \tag{A.6}$$

for

$$\alpha_k\,(w_1,w_2) = \sum_{a_1+a_2=k} \alpha_{a_2}\,(f_{a_1})\,,\ f_{a_1}(t) = w_1(t)\alpha_{a_1}\,(0,t,w_2)\,,$$

and

$$t_n\,(w_1,w_2,w_3) = \sum_{k=0}^{\infty} n^{-k}\alpha_k\,(w_1,w_2,w_3) \tag{A.7}$$

for

$$\alpha_k\,(w_1,w_2,w_3) = \sum_{|a|=k} \alpha_{a_1}\,(g_{a_2,a_3})\,,$$
$$g_{a_2,a_3} = w_1(t)\alpha_{a_2}\,(0,t,f_{a_3})\,,\ f_{a_3} = w_2(t)\alpha_{a_3}\,(0,t,w_3)\,,$$
$$\alpha_0\,(w_1,w_2,w_3) = \int_0^1 g_{0,0} = \int_0^1 w_1\int_0^t w_2\int_0^t w_3,$$
$$\alpha_1\,(w_1,w_2,w_3) = \int_0^1 (g_{0,1} + g_{1,0}) + \left[g_{0,0}\right]_0^1/2,$$
$$g_{0,1} = w_1\int_0^t w_2\left[w_3\right]_0^t/2,\ g_{1,0} = w_1w_2\int_0^t w_3/2,$$

and $\alpha_k\,(0,t,w_1)$ *of (A.3).*

https://doi.org/10.1017/S0269964815000042 Published online by Cambridge University Press

The above propositions can be extended to functions on $[0,1]^r$, as stated by Proposition A.4. Again, we take $K = \infty$ for simplicity.

PROPOSITION A.4: *For* $g: [0,1]^r \to \mathcal{R}$, *set*

$$(g)_{r,n} = n^{-r} \sum_{i_1=1}^{n} \cdots \sum_{i_r=1}^{n} g\left(i_1/n, \ldots, i_r/n\right). \tag{A.8}$$

Then, (A.2) implies

$$(g)_{r,n} = \sum_{k=0}^{\infty} \alpha_{r,k}(g) n^{-k},$$

where

$$\begin{aligned}
\alpha_{r,0}(g) &= \int_0^1 \cdots \int_0^1 g\left(t_1, \ldots, t_r\right)\, dt_1 \cdots dt_r, \\
\alpha_{r,1}(g) &= \sum_{i=1}^{r} \left\{g_i(1) - g_i(0)\right\}/2, \\
g_i\left(t_i\right) &= \left(\int_0^1\right)^{r-1} g\left(t_1, \ldots, t_r\right) dt_1 \cdots dt_{i-1} dt_{i+1} \cdots dt_r, \\
\alpha_{r,n}(g) &= \prod_{i=1}^{r} \left(\sum_{k=0}^{\infty} n^{-k} \beta_{i,k}\right) g\left(t_1, \ldots, t_r\right),
\end{aligned}$$

where the operator $\beta_{i,k}$ *is defined by* $\beta_{i,k} g\left(t_1, \ldots, t_r\right) = \alpha_k(h)$ *for* $h\left(t_i\right) = g\left(t_1, \ldots, t_r\right)$.

For example,

$$\beta_{i,0} g\left(t_1, \ldots, t_r\right) = \int_0^1 g\left(t_1, \ldots, t_r\right)\, dt_i.$$

So,

$$\alpha_{r,k}(g) = \sum \left\{\beta_{1,k_1} \cdots \beta_{r,k_r} g\left(t_1, \ldots, t_r\right) : k_1 + \cdots + k_r = k, k_i \geq 0\right\}.$$

https://doi.org/10.1017/S0269964815000042 Published online by Cambridge University Press