
Department of Applied Mathematics
Faculty of EEMCS



University of Twente
The Netherlands

P.O. Box 217
7500 AE Enschede
The Netherlands

Phone: +31-53-4893400

Fax: +31-53-4893114

Email: memo@math.utwente.nl
www.math.utwente.nl/publications

Memorandum No. 1740

**Tail behavior of the
empirical distribution function
of convolutions**

W. ALBERS AND W.C.M. KALLENBERG

December, 2004

ISSN 0169-2690

Tail behavior of the empirical distribution function of convolutions

Willem Albers and Wilbert C.M. Kallenberg
Department of Applied Mathematics
Faculty of Electrical Engineering, Mathematics and Computer Science
University of Twente
P.O. Box 217, 7500 AE Enschede
The Netherlands

Abstract Control charts based on convolutions require study of the tail behavior of the empirical distribution function of convolutions. It is well-known that this empirical distribution function at a fixed argument x is asymptotically normal. The asymptotic normality is extended here to sequences x_n tending to infinity at a suitable rate. At still larger x_n 's Poisson limiting distributions come in for the classical empirical distribution function. Surprisingly, this property does not generalize to its convolution counterpart, since for those x_n 's it is degenerate at 0 with probability tending to 1. Exact inequalities for the tail behavior are presented as well.

Keyword and phrases: control charts, exceedance probability, convolution sample quantiles, empirical distribution function, U -statistics, Berry-Esseen bound, tail dependence, asymptotic degeneration, Markov, Chebyshev and Bernstein inequalities.

2000 Mathematics Subject Classification: 62E20, 62G30, 62P30

1 Introduction

The mean of a production process variable can be controlled by an individual control chart where an alarm is produced when a new observation exceeds some upper or some lower control limit. The standard Shewhart \bar{X} -chart based on the assumption of normality of the observations takes as control limits $\mu \pm 3\sigma$ with μ and σ the normal mean and standard deviation involved. The probability p that the limits are exceeded during in-control (when indeed the observations come from a $N(\mu, \sigma^2)$ -distribution) is called the false alarm rate and equals here 0.0027. For convenience, in the following we focus on the one-sided case with an upper control limit; the two-sided situation can be treated in a similar way.

When normality fails (as often is the case in practice) the actual false alarm rate may be completely different from the prescribed one, see e.g. Albers et al. (2004). In order to get a false alarm rate p , the upper control limit should in general be taken as the $(1-p)^{th}$ -quantile of the in-control distribution of the observations. In formula, an alarm is produced when $X > F^{-1}(1-p)$, where the random variable X denotes the new observation and F its (continuous) distribution function (df), when the process is in-control. Typically, p is very small and hence, when F is completely unknown, estimation is hardly possible on the basis of earlier, so called Phase I observations when they are not very numerous. For instance, for $p = 0.001$ a number like 200 for the number of Phase I observations is (far) too small to seriously estimate $F^{-1}(1-p)$ in a nonparametric way.

A solution might be to switch to grouped observations with group size k , say, and to estimate the convolution quantile involved by the corresponding sample quantile. To see the advantages (or disadvantages) of such an approach we have to study the behavior of the sample quantiles of convolutions. In fact we come across two extensions compared to classical results about sample quantiles. Firstly, we deal with *convolution* sample quantiles and secondly, we do not consider (only) fixed p , but p_n tending to 0, thus expressing that small values of p are particularly of interest. The first extension is treated in Choudhury and Serfling (1988), see also Arcones (1996). The extension to p_n 's tending to 0 seems not yet to have been done and is one of the main topics of the present paper.

The df of a convolution sample quantile can be linked up with the empirical df of the convolution (see Lemma 1 in Section 2), which in turn can be considered as a so called empirical df of U -statistical structure, see e.g. Silverman (1983). For fixed p asymptotic normality of convolution quantiles holds, see (2.5) in Choudhury and Serfling (1988). We extend the asymptotic normality to p_n 's tending to 0 at a suitable rate. It is well-known that for classical sample quantiles asymptotic normality remains true provided that np_n tends to infinity. On the one hand, when taking convolutions more terms come in, in the sense that the number of (generalized) order statistics is much larger. Or, in other words, that the number of terms in the empirical df of the convolution is much larger. On the other hand, the terms in that U -statistic are dependent. More terms are in general favorable for asymptotic normality, but dependence has a negative influence. In contrast to the situation with classical sample quantiles, the range of p_n 's for which asymptotic normality is obtained for $k \geq 2$ depends on the underlying df F . For instance, when $k = 2$ and F is the exponential distribution, $np_n/\log n$ should tend to infinity, while $np_n\sqrt{\log n} \rightarrow \infty$ suffices when F is the standard normal df. So, for $k \geq 2$ both smaller and larger ranges of admissible p_n 's than for $k = 1$ occur.

For $k = 1$ the empirical df multiplied by n converges to a Poisson distribution when its expectation np_n tends to a positive constant. Surprisingly, this property does not generalize to $k \geq 2$. Although its expectation converges to a positive constant, the empirical df of the convolution multiplied by $\binom{n}{k}$, as a rule still degenerates in the sense that the probability that the empirical df is exactly equal to 0 tends to 1. Asymptotic degeneration at 0 even may occur for p_n 's for which the expectation of $\binom{n}{k}$ times the empirical df converges to infinity.

In fact, the expectation of $\binom{n}{k}$ times the empirical df is, in contrast to suggestions in litera-

ture, not the natural parameter to look at. It turns out that the results of Silverman and Brown (1978) on Poisson limits for U -statistics cannot be applied here. The asymptotic convergence in probability is fully determined by $\lim_{n \rightarrow \infty} n\overline{F}(t_n/k)$, where $\overline{F} = 1 - F$. We get asymptotic degeneration at 0 when this limit equals 0, convergence to ∞ when this limit equals ∞ , and in between a (non-standard!) limit distribution may occur. For instance, when F corresponds to the exponential distribution this limit distribution is not a Poisson distribution, whatever the value of $\lim_{n \rightarrow \infty} n\overline{F}(t_n/k)$.

Exact inequalities on the distribution of the empirical df of the convolution are obtained by using Markov, Chebyshev and Bernstein inequalities.

The paper is organized as follows. In Section 2 notations and definitions are presented and the link between the df of the sample quantiles of the convolution and that of the empirical df of the convolution is obtained. Section 3 is devoted to asymptotic normality. The main tool is a Berry-Esseen bound given in Friedrich (1989). In this way sharper results are derived than when we should apply convergence results in weighted sup-norm metrics for the empirical df of U -statistical structure, cf. Silverman (1983). An asymptotic measure of tail dependence introduced by Embrechts et al. (2002) is refined. The results are exemplified by working out them in detail for the normal and exponential distribution. Section 4 concerns the convergence of $\binom{n}{k}$ times the empirical df as described above. Again the results are exemplified by the normal and exponential distribution. In the final section Markov, Chebyshev and Bernstein inequalities are presented.

2 Notations, definitions and preliminaries

Let X_1, \dots, X_n be i.i.d. r.v.'s with continuous df F . The df of the convolution $X_1 + \dots + X_k$ is denoted by F_k . It is assumed that k is fixed and, keeping the applications to control charts in mind, typically it is rather small, like 2, 3, 4, 5. Write $\overline{F}_k = 1 - F_k$. Define the empirical df of the convolution based on the Phase I observations X_1, \dots, X_n by

$$F_{kn}(x) = \frac{1}{\binom{n}{k}} \sum_{1 \leq i_1 < \dots < i_k \leq n} 1(X_{i_1} + \dots + X_{i_k} \leq x), \quad (1)$$

where $1(A) = 1$ if A holds and 0 otherwise. Note that F_{kn} can also be considered as the empirical df of U -statistical structure, cf. e.g. Silverman (1983), with kernel $h(x_1, \dots, x_k) = x_1 + \dots + x_k$.

For a df G its inverse G^{-1} is defined by

$$G^{-1}(t) = \inf \{x : G(x) \geq t\}$$

and we write

$$\overline{G}^{-1}(t) = G^{-1}(1 - t).$$

The p^{th} upper quantile and p^{th} upper sample quantile of the convolution are defined by

$$\xi_{kp} = \overline{F}_k^{-1}(p), \hat{\xi}_{kp} = \overline{F}_{kn}^{-1}(p),$$

respectively. When ξ_{kp} is known, an alarm is given for Phase II observations X_{n+1}, \dots, X_{n+k} when $X_{n+1} + \dots + X_{n+k} > \xi_{kp}$, yielding a false alarm rate equal to the prescribed p . Since in general ξ_{kp} is unknown we estimate it by $\hat{\xi}_{kp}$. Simply plugging in this estimator leads to an estimation error. Therefore, we apply a correction to it: we take $\hat{\xi}_{kq}$ for some suitable q , to be determined later on. The prescribed false alarm rate will depend on n and therefore we write for it p_n . Similarly, we write q_n in case of the corrected sample quantile. This gives the following procedure. An out-of-control signal is given when

$$X_{n+1} + \dots + X_{n+k} > \hat{\xi}_{kq_n}.$$

Hence, the (conditional) false alarm rate (given $\widehat{\xi}_{kq_n}$) equals

$$P_n = P_n(q_n) = P\left(X_{n+1} + \dots + X_{n+k} > \widehat{\xi}_{kq_n}\right) = \overline{F}_k\left(\widehat{\xi}_{kq_n}\right).$$

Note that the (conditional) false alarm rate P_n is no longer a number, but a r.v. We want to study the distribution of P_n .

Because F is continuous, all $\binom{n}{k}$ values of $X_{i_1} + \dots + X_{i_k}$ with $1 \leq i_1 < \dots < i_k \leq n$ are different w.p.1. We denote them by $X_{(j)}^{(k)}$ for $j = 1, \dots, \binom{n}{k}$ with $X_{(1)}^{(k)} < X_{(2)}^{(k)} < \dots$. Then we obtain

$$\widehat{\xi}_{kq} = X_{\left(\binom{n}{k}\right) - \lfloor \binom{n}{k} q \rfloor}^{(k)}$$

with $\lfloor x \rfloor$ being the entier of x . The following result entails an expression for the df of P_n , thus linking the df of the sample quantile of the convolution with the df of F_{kn} , the empirical df of the convolution.

Lemma 1 For all $0 < \eta < 1$ and $0 < q < 1$

$$P(P_n(q) > \eta) = P\left(\overline{F}_{kn}\left(\overline{F}_k^{-1}(\eta)\right) \leq \frac{\lfloor \binom{n}{k} q \rfloor}{\binom{n}{k}}\right).$$

Proof. Since for each df G it holds that

$$G(x) < \eta \iff x < G^{-1}(\eta)$$

and hence

$$\overline{G}(x) > \eta \iff x < \overline{G}^{-1}(\eta),$$

we get for all $0 < \eta < 1$ and $0 < q < 1$

$$P(P_n(q) > \eta) = P\left(\overline{F}_k\left(\widehat{\xi}_{kq}\right) > \eta\right) = P\left(\widehat{\xi}_{kq} < \overline{F}_k^{-1}(\eta)\right).$$

Moreover, we have w.p.1, for all $x \in \mathbb{R}$,

$$\begin{aligned} X_{(j)}^{(k)} < x &\iff X_{(j)}^{(k)} \leq x \iff \frac{j}{\binom{n}{k}} = F_{kn}\left(X_{(j)}^{(k)}\right) \leq F_{kn}(x) \\ &\iff 1 - \frac{j}{\binom{n}{k}} \geq \overline{F}_{kn}(x) \end{aligned}$$

and therefore

$$\begin{aligned} P(P_n(q) > \eta) &= P\left(\widehat{\xi}_{kq} < \overline{F}_k^{-1}(\eta)\right) = P\left(X_{\left(\binom{n}{k}\right) - \lfloor \binom{n}{k} q \rfloor}^{(k)} < \overline{F}_k^{-1}(\eta)\right) \\ &= P\left(\overline{F}_{kn}\left(\overline{F}_k^{-1}(\eta)\right) \leq \frac{\lfloor \binom{n}{k} q \rfloor}{\binom{n}{k}}\right), \end{aligned}$$

as was to be proved. ■

We are interested in the asymptotic distribution of $P_n(q_n)$ with q_n tending to 0 as $n \rightarrow \infty$, while the interesting values of η are $\eta = \eta_n = p_n(1 + \varepsilon_n)$ for suitable $\varepsilon_n \geq 0$. In particular, we want to have, for some (fixed) $0 < \alpha < 1$,

$$\lim_{n \rightarrow \infty} P(P_n(q_n) > p_n(1 + \varepsilon_n)) = \alpha.$$

This is the so called exceedance probability approach. It ensures that the probability with which values of P_n occur that are too unpleasant (that is more than a factor $1 + \varepsilon_n$ larger than the

prescribed false alarm rate p_n) is not large ($\alpha = 0.1$, say). For a more detailed discussion on this criterion, see Albers and Kallenberg (2004a, b), Albers et al. (2002).

For obtaining $\lim_{n \rightarrow \infty} P(P_n(q_n) > p_n(1 + \varepsilon_n)) = \alpha$, we can in view of Lemma 1 equivalently require

$$\lim_{n \rightarrow \infty} P\left(\overline{F}_{kn}\left(\overline{F}_k^{-1}(p_n(1 + \varepsilon_n))\right) \leq \tilde{q}_n\right) = \alpha, \text{ with } \tilde{q}_n = \frac{\binom{n}{k} q_n}{\binom{n}{k}}. \quad (2)$$

We study the asymptotic distribution of $\overline{F}_{kn}(t_n)$ and apply this for $t_n = \overline{F}_k^{-1}(p_n(1 + \varepsilon_n))$.

Remark 2.1 Consider $k = 1$. We then have $P_n(q_n) = \overline{F}_1(\hat{\xi}_{1q_n}) = \overline{F}(X_{(n-r_n)})$ with $r_n = [nq_n]$. The distribution of $\overline{F}(X_{(n-r_n)})$ is the same as that of $U_{(r_n+1)}$, being the $(r_n + 1)^{th}$ order statistic of a sample of uniform r.v.'s. A similar (simple) connection with uniform distributions does not seem to exist for $k \geq 2$.

3 Asymptotic normality

From its definition, see (1) it is seen that $\overline{F}_{kn}(t_n)$ is a U -statistic with a kernel depending on n . We introduce the following notation

$$\{s_n(t)\}^2 = P(X_1 + X_2 + \dots + X_k > t, X_1 + X_2^* + \dots + X_k^* > t) - \{\overline{F}_k(t)\}^2, \quad (3)$$

where $X_1, X_2, \dots, X_k, X_2^*, \dots, X_k^*$ are i.i.d. random variables with df F . Furthermore, let

$$h_n(X_1, \dots, X_k; t_n) = \frac{1(X_1 + \dots + X_k > t_n) - \overline{F}_k(t_n)}{s_n(t_n)}$$

and denote its projection by

$$\tilde{h}_n(x_j; t_n) = E(h_n(X_1, \dots, X_k; t_n) | X_j = x_j) = \frac{\overline{F}_{k-1}(t_n - x_j) - \overline{F}_k(t_n)}{s_n(t_n)},$$

then

$$E\tilde{h}_n(X_j; t_n) = 0, \text{ var}\left(\tilde{h}_n(X_j; t_n)\right) = 1. \quad (4)$$

Writing

$$T_n(t_n) = \frac{1}{\binom{n}{k}} \sum_{1 \leq i_1 < \dots < i_k \leq n} h_n(X_{i_1}, \dots, X_{i_k}; t_n),$$

we get

$$T_n(t_n) = \frac{\overline{F}_{kn}(t_n) - \overline{F}_k(t_n)}{s_n(t_n)} \quad (5)$$

and thus

$$P(\overline{F}_{kn}(t_n) \leq \tilde{q}_n) = P\left(T_n(t_n) \leq \frac{\tilde{q}_n - \overline{F}_k(t_n)}{s_n(t_n)}\right).$$

In order to establish asymptotic normality we cannot use (directly) limiting results for U -statistics, since our kernel depends on n . Therefore, we rely on a Berry-Esseen bound given by Friedrich (1989), but see also Remark 3.2. For the sharpness of the Berry-Esseen bound see Bentkus et al. (1994). Our notation is linked up with the notation in Friedrich (1989).

Theorem 2 *Define*

$$\begin{aligned} \gamma_{0,n} &= n^{-1/2} E \left| \frac{\overline{F}_{k-1}(t_n - X_1) - \overline{F}_k(t_n)}{s_n(t_n)} \right|^3, \\ \gamma_{3,r,n} &= \frac{4(k-1)}{n^{1/2}(n-1)} \frac{[\overline{F}_k(t_n) \{1 - \overline{F}_k(t_n)\}]^r + \{1 - \overline{F}_k(t_n)\} [\overline{F}_k(t_n)]^r}{s_n(t_n)} \text{ for } r \geq 1. \end{aligned} \quad (6)$$

Then there exists a constant $C \in \mathbb{R}$, such that for $\frac{3}{2} \leq r < 2$

$$\sup_{x \in \mathbb{R}} \left| P \left(\frac{\sqrt{n}}{k} T_n(t_n) \leq x \right) - \Phi(x) \right| \leq C \left(\gamma_{0,n} + \frac{1}{2-r} n^{13/6} \gamma_{0,n}^{1/3} \gamma_{3,r,n}^r + n^{4/3} \gamma_{0,n}^{2/3} \gamma_{3,3/2,n} \right),$$

where Φ denotes the df of the standard normal distribution. The estimate remains true for $r = 2$ if $1/(2-r)$ is replaced by $\log n$.

Proof. We apply the Berry-Esseen bound given in Theorem 2.1 (a), (c) of Friedrich (1989). We apply it for the case of U -statistics and use the arguments of Remark 5 in Friedrich (1989). We have

$$\begin{aligned} E[T_n(t_n) | X_j] &= \frac{\binom{n-1}{k-1} \{ \bar{F}_{k-1}(t_n - X_j) - \bar{F}_k(t_n) \}}{\binom{n}{k} s_n(t_n)} \\ &= \frac{k \bar{F}_{k-1}(t_n - X_j) - \bar{F}_k(t_n)}{n s_n(t_n)} \\ &= \frac{k}{n} \tilde{h}_n(X_j; t_n). \end{aligned}$$

In view of (4) we obtain

$$E T_n(t_n) = 0 \text{ and } \text{var} \left(\sum_{j=1}^n E[T_n(t_n) | X_j] \right) = \frac{k^2}{n}.$$

It is now immediately seen that γ_0 and γ_1 in Friedrich (1989) boil down to our $\gamma_{0,n}$ and $n^{-1/2} (n^{1/2} \gamma_{0,n})^{1/3} = n^{-1/3} \gamma_{0,n}^{1/3}$, respectively.

Write for $j = 1, \dots, n$,

$$\check{X}_j = (X_1, \dots, X_{j-1}, X_{j+1}, \dots, X_n)$$

and for $1 \leq j < m \leq n$,

$$\begin{aligned} \check{X}_{jm} &= (X_1, \dots, X_{j-1}, X_{j+1}, \dots, X_{m-1}, X_{m+1}, \dots, X_n), \\ \hat{T}_{jm,n}(t_n) &= E[T_n(t_n) | \check{X}_j] + E[T_n(t_n) | \check{X}_m] - E[T_n(t_n) | \check{X}_{jm}], \\ D_{jm}[T_n(t_n)] &= T_n(t_n) - \hat{T}_{jm,n}(t_n). \end{aligned}$$

Then

$$D_{jm}[T_n(t_n)] = \frac{1}{\binom{n}{k}} \sum_{1 \leq i_1 < \dots < i_k \leq n} \sum_{j, m \in \{i_1, \dots, i_k\}} \tilde{h}_{jm,n}(X_{i_1}, \dots, X_{i_k}; t_n),$$

where, writing E_j for taking expectation w.r.t. X_j and E_{jm} for taking expectation w.r.t. X_j and X_m ,

$$\begin{aligned} \tilde{h}_{jm,n}(X_{i_1}, \dots, X_{i_k}; t_n) &= h_n(X_{i_1}, \dots, X_{i_k}; t_n) - E_j h_n(X_{i_1}, \dots, X_{i_k}; t_n) \\ &\quad - E_m h_n(X_{i_1}, \dots, X_{i_k}; t_n) + E_{jm} h_n(X_{i_1}, \dots, X_{i_k}; t_n). \end{aligned}$$

Employing Minkowski's inequality and inequalities (for $r \geq 1$) like

$$|E_j h_n(X_{i_1}, \dots, X_{i_k}; t_n)|^r \leq \{E_j |h_n(X_{i_1}, \dots, X_{i_k}; t_n)|\}^r \leq E_j |h_n(X_{i_1}, \dots, X_{i_k}; t_n)|^r,$$

we get (for $r \geq 1$)

$$\begin{aligned} \{E |D_{jm}[T_n(t_n)]|^r\}^{1/r} &\leq \frac{4 \binom{n-2}{k-2}}{\binom{n}{k}} \{E |h_n(X_1, \dots, X_k; t_n)|^r\}^{1/r} \\ &= \frac{4k(k-1)}{n(n-1)} \frac{[\bar{F}_k(t_n) \{1 - \bar{F}_k(t_n)\}^r + \{1 - \bar{F}_k(t_n)\} \{\bar{F}_k(t_n)\}^r]^{1/r}}{s_n(t_n)}. \end{aligned}$$

Application of Theorem 2.1. in Friedrich (1989) completes the proof. ■

In Embrechts et al. (2002) an asymptotic measure of tail dependence is introduced as follows. Let Y_1 and Y_2 be r.v.'s with df's G_1 and G_2 . The coefficient of upper tail dependence of Y_1 and Y_2 is $\lim_{\eta \downarrow 0} P\left(Y_2 > \overline{G}_2^{-1}(\eta) \mid Y_1 > \overline{G}_1^{-1}(\eta)\right)$. Here we are interested in the tail behavior of such conditional probabilities as well. Let $S_k = X_1 + X_2 + \dots + X_k$, $S_k^* = X_1 + X_2^* + \dots + X_k^*$ and $\eta_n = \overline{F}_k(t_n)$. We may write, see (3),

$$\begin{aligned} \{s_n(t_n)\}^2 &= P(X_1 + X_2 + \dots + X_k > t_n, X_1 + X_2^* + \dots + X_k^* > t_n) - \overline{F}_k(t_n)^2 \\ &= P\left(S_k > \overline{F}_k^{-1}(\eta_n) \mid S_k^* > \overline{F}_k^{-1}(\eta_n)\right) \eta_n - \eta_n^2. \end{aligned}$$

When the coefficient of upper tail dependence of S_k and S_k^* equals $\zeta > 0$, S_k and S_k^* are called asymptotically dependent and $\{s_n(t_n)\}^2 = \zeta \eta_n(1 + o(1))$. Here we often come across asymptotic independence and a more refined analysis is needed, taking into account the rate at which $P\left(S_k > \overline{F}_k^{-1}(\eta_n) \mid S_k^* > \overline{F}_k^{-1}(\eta_n)\right)$ tends to 0. If $P\left(S_k > \overline{F}_k^{-1}(\eta_n) \mid S_k^* > \overline{F}_k^{-1}(\eta_n)\right)$ is of order $g(\eta_n)$ for some function g , we call g the upper tail dependence rate of S_k and S_k^* . More precisely, g is the upper tail dependence rate of S_k and S_k^* if

$$0 < \lim_{n \rightarrow \infty} \frac{P\left(S_k > \overline{F}_k^{-1}(\eta_n) \mid S_k^* > \overline{F}_k^{-1}(\eta_n)\right)}{g(\eta_n)} < \infty.$$

Example 3.1 Let $k = 2$ and consider for F the standard normal df Φ . Let (X, Y) have a standard bivariate normal distribution with correlation coefficient ρ . Then application of well-known results on Mill's ratio gives for $t \rightarrow \infty$, $-1 < \rho \leq 1$,

$$P(Y > t \mid X > t) = (1 + \rho) \overline{\Phi}\left(t \sqrt{\frac{1 - \rho}{1 + \rho}}\right) (1 + o(1)). \quad (7)$$

A detailed proof of (7) when $-1 < \rho < 1$ is given in the Appendix. For $\rho = 1$, (7) obviously is true. Interpreting the right-hand side of (7) as 0, the result continuous to hold for $\rho = -1$ as well. (Note that on page 199 in Embrechts et al. (2002) it is incorrectly suggested that the factor $(1 + \rho)$ in (7) should be 2.) Moreover, writing φ for the standard normal density, we obtain

$$\eta_n = \overline{F}_2(t_n) = \overline{\Phi}\left(\frac{t_n}{\sqrt{2}}\right) = \frac{\varphi(2^{-1/2}t_n)}{2^{-1/2}t_n} (1 + o(1)).$$

Hence, we get with $g(\eta) = \eta^{1/3} |\log \eta|^{-1/3}$

$$\begin{aligned} & \lim_{n \rightarrow \infty} \frac{P\left(S_2 > \overline{F}_2^{-1}(\eta_n) \mid S_2^* > \overline{F}_2^{-1}(\eta_n)\right)}{g(\eta_n)} \\ &= \lim_{n \rightarrow \infty} \frac{P(S_2 > t_n \mid S_2^* > t_n)}{g(\eta_n)} \\ &= \lim_{n \rightarrow \infty} \frac{\frac{3}{2} \overline{\Phi}\left(\frac{t_n}{\sqrt{6}}\right)}{g(\eta_n)} = 3^{3/2} 2^{-5/3} \pi^{-1/3}. \end{aligned}$$

Therefore, in case of normality, $g(\eta) = \eta^{1/3} |\log \eta|^{-1/3}$ is the upper tail dependence rate of S_2 and S_2^* .

Example 3.2 Let $k = 2$ and consider for F the exponential df with parameter $\lambda = 1$. Direct calculation gives

$$\begin{aligned} \eta_n &= \overline{F}_2(t_n) = (1 + t_n) e^{-t_n}, \\ P(S_2 > t_n, S_2^* > t_n) &= 2e^{-t_n} - e^{-2t_n} \end{aligned}$$

and hence

$$\begin{aligned} & \lim_{n \rightarrow \infty} P \left(S_2 > \overline{F}_2^{-1}(\eta_n) | S_2^* > \overline{F}_2^{-1}(\eta_n) \right) |\log \eta_n| \\ &= \lim_{n \rightarrow \infty} P \left(S_2 > t_n | S_2^* > t_n \right) t_n = 2. \end{aligned}$$

We conclude that for the exponential distribution $g(\eta) = |\log \eta|^{-1}$ is the upper tail dependence rate of S_2 and S_2^* .

Corollary 3 *Let $\eta_n = \overline{F}_k(t_n)$ and let g be the upper tail dependence rate of S_k and S_k^* . Suppose that*

$$E \left| \overline{F}_{k-1}(t_n - X_1) - \overline{F}_k(t_n) \right|^3 = c^* g^*(\eta_n) (1 + o(1)) \text{ for some } c^* > 0. \quad (8)$$

Assume that

$$\begin{aligned} \frac{\eta_n}{g(\eta_n)} &= o(1), \\ \eta_n \{ \eta_n g(\eta_n) \}^{1/6} \{ g^*(\eta_n) \}^{-2/3} &= O(1), \\ \eta_n^2 \{ g^*(\eta_n) \}^{-1} &= O(1). \end{aligned} \quad (9)$$

Then

$$\sup_{x \in \mathbb{R}} \left| P \left(\frac{\sqrt{n}}{k} T_n(t_n) \leq x \right) - \Phi(x) \right| \rightarrow 0$$

for those t_n such that

$$n^{-1/2} g^*(\eta_n) \{ \eta_n g(\eta_n) \}^{-3/2} \rightarrow 0.$$

Proof. Since g is the upper tail dependence rate of S_k and S_k^* and $\eta_n/g(\eta_n) = o(1)$ implying $\eta_n^2 = o(\eta_n g(\eta_n))$, we have

$$\{s_n(t_n)\}^2 = c \eta_n g(\eta_n) (1 + o(1)) \text{ for some } c > 0.$$

Hence, in view of (6), we get

$$\begin{aligned} \gamma_{0,n} &= n^{-1/2} c^* g^*(\eta_n) \{ c \eta_n g(\eta_n) \}^{-3/2} (1 + o(1)), \\ \gamma_{3,r,n} &= \tilde{c} n^{-3/2} \eta_n^{1/r} \{ \eta_n g(\eta_n) \}^{-1/2} (1 + o(1)) \end{aligned}$$

for some $\tilde{c} > 0$. Using (9) we easily obtain

$$\begin{aligned} n^{13/6} \gamma_{0,n}^{1/3} \gamma_{3,5/3,n}^{5/3} &= O(\gamma_{0,n}), \\ n^{4/3} \gamma_{0,n}^{2/3} \gamma_{3,3/2,n} &= O(\gamma_{0,n}) \end{aligned}$$

and thus, by Theorem 2,

$$\sup_{x \in \mathbb{R}} \left| P \left(\frac{\sqrt{n}}{k} T_n(t_n) \leq x \right) - \Phi(x) \right| = O(\gamma_{0,n}).$$

Because $\gamma_{0,n} \rightarrow 0$ iff $n^{-1/2} g^*(\eta_n) \{ \eta_n g(\eta_n) \}^{-3/2} \rightarrow 0$, the proof is complete. ■

Remark 3.1 It is easily checked that Corollary 3 continuous to hold if (8) is replaced by

$$0 < \liminf_{n \rightarrow \infty} \frac{E \left| \overline{F}_{k-1}(t_n - X_1) - \overline{F}_k(t_n) \right|^3}{g^*(\eta_n)} \leq \limsup_{n \rightarrow \infty} \frac{E \left| \overline{F}_{k-1}(t_n - X_1) - \overline{F}_k(t_n) \right|^3}{g^*(\eta_n)} < \infty. \quad (10)$$

Similarly, the condition that g is the upper tail dependence rate of S_k and S_k^* may be replaced by

$$0 < \liminf_{n \rightarrow \infty} \frac{P\left(S_k > \bar{F}_k^{-1}(\eta_n) \mid S_k^* > \bar{F}_k^{-1}(\eta_n)\right)}{g(\eta_n)} \leq \limsup_{n \rightarrow \infty} \frac{P\left(S_k > \bar{F}_k^{-1}(\eta_n) \mid S_k^* > \bar{F}_k^{-1}(\eta_n)\right)}{g(\eta_n)} < \infty.$$

In order to get an idea for which p_n 's we get asymptotic normality we consider the following examples.

Example 3.3 Let $k = 2$ and consider for F the standard normal df Φ . From Example 3.1 we know that $g(\eta) = \eta^{1/3} |\log \eta|^{-1/3}$ is the upper tail dependence rate of S_2 and S_2^* . Next it will be shown that (10) holds with

$$g^*(\eta_n) = \eta_n^{3/2} |\log(\eta_n)|^{-3/4}.$$

Let $\delta > 0$ be a sufficiently small constant. Then we get for some constant $c_1 > 0$ (noting that the $o(1)$ -terms occurring in the integrals do not depend on x)

$$\begin{aligned} & E \left| \bar{F}_{k-1}(t_n - X_1) - \bar{F}_k(t_n) \right|^3 \\ &= E \left| \bar{\Phi}(t_n - X_1) - \bar{\Phi}(2^{-1/2}t_n) \right|^3 \\ &\geq \int_{(1-\delta)3t_n/4}^{(1+\delta)3t_n/4} \left\{ \bar{\Phi}(t_n - x) \right\}^3 (1 + o(1)) \varphi(x) dx \\ &= \int_{(1-\delta)3t_n/4}^{(1+\delta)3t_n/4} \left\{ \frac{\varphi(t_n - x)}{t_n - x} \right\}^3 (1 + o(1)) \varphi(x) dx \\ &\geq c_1 t_n^{-3} e^{-3t_n^2/8} \int_{(1-\delta)3t_n/4}^{(1+\delta)3t_n/4} \varphi(2x - 3t_n/2) dx \\ &= c_1 t_n^{-3} e^{-3t_n^2/8} (1 + o(1)). \end{aligned}$$

On the other hand, we have

$$\begin{aligned} & E \left| \bar{\Phi}(t_n - X_1) - \bar{\Phi}(2^{-1/2}t_n) \right|^3 \\ &\leq E \left\{ \bar{\Phi}(t_n - X_1) \right\}^3 + \left\{ \bar{\Phi}(2^{-1/2}t_n) \right\}^3 \\ &= E \left\{ \bar{\Phi}(t_n - X_1) \right\}^3 + O\left(t_n^{-3} e^{-3t_n^2/4}\right) \end{aligned}$$

and, for some constant $c_2 > 0$,

$$\begin{aligned} E \left\{ \bar{\Phi}(t_n - X_1) \right\}^3 &\leq \int_{-\infty}^{(1-\delta)t_n} \left\{ \bar{\Phi}(t_n - x) \right\}^3 \varphi(x) dx + \bar{\Phi}((1-\delta)t_n) \\ &\leq (\delta t_n)^{-3} \int_{-\infty}^{(1-\delta)t_n} \left\{ \varphi(t_n - x) \right\}^3 \varphi(x) dx + O\left(t_n^{-1} e^{-(1-\delta)t_n^2/2}\right) \\ &\leq c_2 t_n^{-3} e^{-3t_n^2/8} \int_{-\infty}^{(1-\delta)t_n} \varphi(2x - 3t_n/2) dx + O\left(t_n^{-1} e^{-(1-\delta)t_n^2/2}\right). \end{aligned}$$

Noting that for some constant $c_3 > 0$,

$$t_n^{-3} e^{-3t_n^2/8} = c_3 \eta_n^{3/2} |\log(\eta_n)|^{-3/4} (1 + o(1)),$$

the result follows.

Since $\eta_n |\log(\eta_n)| \rightarrow 0$, it easily follows that the conditions of Corollary 3 hold and hence asymptotic normality is obtained for those t_n such that

$$n^{-1/2} g^*(\eta_n) \{\eta_n g(\eta_n)\}^{-3/2} \rightarrow 0.$$

That is when

$$n\eta_n |\log(\eta_n)|^{1/2} \rightarrow \infty.$$

Taking $\eta_n = p_n(1 + \varepsilon_n)$ with ε_n bounded, asymptotic normality holds when

$$\lim_{n \rightarrow \infty} np_n |\log p_n|^{1/2} = \infty \text{ or } p_n = \frac{a_n}{n\sqrt{\log n}} \text{ with } a_n \rightarrow \infty.$$

Example 3.4 Let $k = 2$ and consider the exponential distribution with parameter $\lambda = 1$. From Example 3.2 we know that $g(\eta) = |\log \eta|^{-1}$ is the upper tail dependence rate of S_2 and S_2^* . Direct calculation gives

$$g^*(\eta_n) = \eta_n |\log(\eta_n)|^{-1}.$$

Since $\eta_n |\log(\eta_n)| \rightarrow 0$, it easily follows that the conditions of Corollary 3 hold and hence asymptotic normality is obtained for those t_n such that

$$n^{-1/2} g^*(\eta_n) \{\eta_n g(\eta_n)\}^{-3/2} \rightarrow 0.$$

That is when

$$n\eta_n |\log(\eta_n)|^{-1} \rightarrow \infty.$$

Taking $\eta_n = p_n(1 + \varepsilon_n)$ with ε_n bounded, asymptotic normality holds when

$$\lim_{n \rightarrow \infty} \frac{np_n}{|\log p_n|} = \infty \text{ or } p_n = \frac{a_n \log n}{n} \text{ with } a_n \rightarrow \infty.$$

It is seen from Examples 3.3 and 3.4 that compared to $k = 1$, where asymptotic normality is obtained when $\lim_{n \rightarrow \infty} np_n = \infty$, a relaxation in the sense of a larger range of p_n 's for which asymptotic normality holds is possible (Example 3.3) as well as a restriction to a smaller ranges of admissible p_n 's (Example 3.4), depending on the df of the observations. Note that this is in contrast to $k = 1$, where the range of p_n 's leading to asymptotic normality does not depend on F (other than through p_n itself).

Remark 3.2 Another potential way to derive asymptotic normality for $\overline{F}_{kn}(t_n)$ is an application of convergence results in weighted sup-norm metrics for the empirical df of U -statistical structure, cf. Silverman (1983). Obviously, convergence of the supremum over t implies convergence at each sequence t_n . Invoking Theorem A of Silverman (1983) gives

$$v(\overline{F}_k(t_n)) \frac{n^{1/2} \{\overline{F}_{kn}(t_n) - \overline{F}_k(t_n)\} - ks_n(t_n)W^*}{\{\overline{F}_k(t_n)\}^{1/2}} \rightarrow 0,$$

where W^* has a standard normal df and $\lim_{t \rightarrow 0} v(t) = 0$. Equivalently, we get

$$\frac{v(\overline{F}_k(t_n)) s_n(t_n)}{\{\overline{F}_k(t_n)\}^{1/2}} \left(\frac{\sqrt{n}}{k} T_n(t_n) - W^* \right) \rightarrow 0. \quad (11)$$

As a rule we have $s_n(t_n) = o(\{\overline{F}_k(t_n)\}^{1/2})$, see also Examples 3.1 and 3.2, and, moreover, $v(t) \rightarrow 0$, implying that (11) in general is a much weaker result than Corollary 3. Indeed, the convergence results in weighted sup-norm metrics for the empirical df of U -statistical structure are less tailored to our situation here than the Berry-Esseen bound.

Next we apply the asymptotic normality to get asymptotic exceedance probability equal to α , cf. (2).

Theorem 4 Let $\eta_n = p_n(1 + \varepsilon_n) = \bar{F}_k(t_n)$. Suppose that for some $c > 0$

$$\lim_{n \rightarrow \infty} \frac{P\left(S_k > \bar{F}_k^{-1}(\eta_n) \mid S_k^* > \bar{F}_k^{-1}(\eta_n)\right)}{g(\eta_n)} = c,$$

$$0 < \liminf_{n \rightarrow \infty} \frac{E\left|\bar{F}_{k-1}(t_n - X_1) - \bar{F}_k(t_n)\right|^3}{g^*(\eta_n)} \leq \limsup_{n \rightarrow \infty} \frac{E\left|\bar{F}_{k-1}(t_n - X_1) - \bar{F}_k(t_n)\right|^3}{g^*(\eta_n)} < \infty.$$

Assume that

$$\begin{aligned} \frac{\eta_n}{g(\eta_n)} &= o(1), \\ \eta_n (\eta_n g(\eta_n))^{1/6} \{g^*(\eta_n)\}^{-2/3} &= O(1), \\ \eta_n^2 \{g^*(\eta_n)\}^{-1} &= O(1), \\ n^{-1/2} g^*(\eta_n) \{\eta_n g(\eta_n)\}^{-3/2} &= o(1), \\ n^{2k-1} \eta_n g(\eta_n) &\rightarrow \infty. \end{aligned}$$

Let

$$q_n = p_n(1 + \varepsilon_n) - kn^{-1/2} \{c\eta_n g(\eta_n)\}^{1/2} \bar{\Phi}^{-1}(\alpha), \quad (12)$$

then

$$\lim_{n \rightarrow \infty} P(P_n(q_n) > p_n(1 + \varepsilon_n)) = \alpha.$$

Proof. Application of Corollary 3 (see also Remark 3.1) gives

$$\sup_{x \in \mathbb{R}} \left| P\left(\frac{\sqrt{n}}{k} T_n(t_n) \leq x\right) - \Phi(x) \right| \rightarrow 0,$$

where, in view of (5),

$$T_n(t_n) = \frac{\bar{F}_{kn}(t_n) - \bar{F}_k(t_n)}{s_n(t_n)}.$$

Hence, by Lemma 1, see also (2),

$$\begin{aligned} &\lim_{n \rightarrow \infty} P(P_n(q_n) > p_n(1 + \varepsilon_n)) \\ &= \lim_{n \rightarrow \infty} P\left(\bar{F}_{kn}\left(\bar{F}_k^{-1}(p_n(1 + \varepsilon_n))\right) \leq \tilde{q}_n\right) \\ &= \lim_{n \rightarrow \infty} P\left(\bar{F}_{kn}(t_n) \leq \tilde{q}_n\right) \\ &= \lim_{n \rightarrow \infty} P\left(T_n(t_n) \leq \frac{\tilde{q}_n - p_n(1 + \varepsilon_n)}{s_n(t_n)}\right) \\ &= \lim_{n \rightarrow \infty} \Phi\left(\frac{\sqrt{n}}{k} \frac{\tilde{q}_n - p_n(1 + \varepsilon_n)}{s_n(t_n)}\right) = \alpha, \end{aligned}$$

since, using $-\bar{\Phi}^{-1}(\alpha) = \Phi^{-1}(\alpha)$,

$$\begin{aligned} &\frac{\sqrt{n}}{k} \frac{\tilde{q}_n - p_n(1 + \varepsilon_n)}{s_n(t_n)} \\ &= \frac{\sqrt{n}}{k} \frac{kn^{-1/2} \{c\eta_n g(\eta_n)\}^{1/2} \bar{\Phi}^{-1}(\alpha) + O(n^{-k})}{\{c\eta_n g(\eta_n)\}^{1/2} (1 + o(1))} \\ &= \bar{\Phi}^{-1}(\alpha) (1 + o(1)) + O\left(n^{-k+1/2} \{\eta_n g(\eta_n)\}^{-1/2}\right) \\ &= \bar{\Phi}^{-1}(\alpha) + o(1). \end{aligned}$$

This completes the proof of the theorem. ■

Remark 3.3 For $k = 1$ we get (assuming only $np_n \rightarrow \infty$)

$$q_n = p_n(1 + \varepsilon_n) - n^{-1/2} \{p_n(1 + \varepsilon_n)\}^{1/2} \overline{\Phi}^{-1}(\alpha),$$

which coincides with (12) when taking $k = 1$ and $c = 1, g \equiv 1$. Note that, given p_n , for $k = 1$ the correction to get the asymptotic exceedance probability equal to α does not depend on the df F , while for $k \geq 2$ it does, since g and c depend on F . As a rule $g(\eta_n)$ tends to 0 and hence the correction term for $k \geq 2$ is smaller than the one for $k = 1$. In fact, the factor $k \{cg(\eta_n)\}^{1/2}$ expresses the gain when taking the convolution chart instead of the individual chart.

Example 3.5 Let $k = 2$ and consider for F the standard normal df Φ . Assume that $\lim_{n \rightarrow \infty} p_n n \sqrt{\log n} = \infty$. The conditions of Theorem 4 are checked in Example 3.3, except for the additional condition $n^{2k-1} \eta_n g(\eta_n) \rightarrow \infty$. (Note that $c = 3^{3/2} 2^{-5/3} \pi^{-1/3}$, see Example 3.1.) It is easily seen that this additional condition also holds. Therefore the asymptotic exceedance probability equals α , when taking

$$q_n = p_n(1 + \varepsilon_n) - 2^{1/6} 3^{3/4} \pi^{-1/6} n^{-1/2} \eta_n^{2/3} |\log \eta_n|^{-1/6} \overline{\Phi}^{-1}(\alpha). \quad (13)$$

For instance, when taking $p_n = n^{-1}$ and $\varepsilon_n = 0$, this yields

$$q_n = p_n \left\{ 1 - 2^{1/6} 3^{3/4} \pi^{-1/6} (n \log n)^{-1/6} \overline{\Phi}^{-1}(\alpha) \right\}.$$

The gain when taking the convolution chart with $k = 2$ instead of the individual chart is a reduction of the correction term with a factor $2^{1/6} 3^{3/4} \pi^{-1/6} \eta_n^{1/6} |\log \eta_n|^{-1/6}$.

Example 3.6 Let $k = 2$ and consider the exponential distribution with parameter $\lambda = 1$. Assume that $\lim_{n \rightarrow \infty} p_n n (\log n)^{-1} = \infty$. The conditions of Theorem 4 are checked in Example 3.4, except for the additional condition $n^{2k-1} \eta_n g(\eta_n) \rightarrow \infty$. (Note that $c = 2$, see Example 3.2.) It is easily seen that this additional condition also holds. Therefore, the asymptotic exceedance probability equals α , when taking

$$q_n = p_n(1 + \varepsilon_n) - 2^{3/2} n^{-1/2} \eta_n^{1/2} |\log \eta_n|^{-1/2} \overline{\Phi}^{-1}(\alpha). \quad (14)$$

For instance, when taking $p_n = n^{-1} (\log n)^2$ and $\varepsilon_n = 0$, we may take

$$q_n = p_n \left\{ 1 - 2^{3/2} (\log n)^{-3/2} \overline{\Phi}^{-1}(\alpha) \right\}.$$

The gain when taking the convolution chart with $k = 2$ instead of the individual chart is a reduction of the correction term with a factor $2^{3/2} |\log \eta_n|^{-1/2}$.

Remark 3.4 Some improvement can be made in the theorems and the corollary of this section by the following argument. Asymptotic normality of a sum of r.v.'s may be destroyed when a few terms are much larger than the others, see the well-known Lindeberg-Feller condition. Here this may occur through some large value of one X_i , giving a lot of indicator-terms in $\overline{F}_{kn}(t_n)$ the value 1. On the other hand, the probability of getting such a large value is small. In particular, let u_n satisfy

$$\lim_{n \rightarrow \infty} n \overline{F}(u_n) = 0,$$

then

$$\lim_{n \rightarrow \infty} P \left(\max_{1 \leq i \leq n} X_i \leq u_n \right) = 1.$$

Hence,

$$\begin{aligned} P\left(\frac{\sqrt{n}}{k}T_n(t_n) \leq x\right) &= P\left(\frac{\sqrt{n}}{k}T_n(t_n) \leq x, \max_{1 \leq i \leq n} X_i \leq u_n\right) + o(1) \\ &= P\left(\frac{\sqrt{n}}{k}T_n(t_n) \leq x | X_i \leq u_n \text{ for all } 1 \leq i \leq n\right) + o(1) \end{aligned}$$

and we may replace in all the conditions F by the df of X given $X \leq u_n$. We shall come back to this u_n in the next section (see Remark 4.3).

4 Convergence of $\binom{n}{k}F_{kn}$

The empirical df of X_1, \dots, X_n , defined by $F_n(t) = n^{-1} \sum_{i=1}^n 1(X_i \leq t)$, has a limiting distribution that is completely elaborated: when $nF(t_n)\overline{F}(t_n) \rightarrow \infty$, we get asymptotic normality for the standardized F_n or \overline{F}_n ; when $n\overline{F}(t_n) \rightarrow 0$, we get that $n\overline{F}_n(t_n)$ converges in probability to 0 and in between, when $n\overline{F}(t_n) \rightarrow c$ the limiting distribution of $n\overline{F}_n(t_n)$ is the Poisson distribution with parameter c and similarly for $nF_n(t_n)$. In the previous section we have seen that asymptotic normality for $\overline{F}_{kn}(t_n)$ can be established under certain conditions on t_n , which may lead to a larger or smaller range of admissible $\overline{F}_k(t_n)$'s than for the empirical df of X_1, \dots, X_n itself.

In this section we will start with investigating a possible Poisson limit for $\binom{n}{k}\overline{F}_{kn}(t_n)$. For getting a Poisson limit distribution one would guess that the following condition should be required

$$\lim_{n \rightarrow \infty} \binom{n}{k} \overline{F}_k(t_n) = c \quad (15)$$

for some $c > 0$, implying

$$\lim_{n \rightarrow \infty} \binom{n}{k} E\overline{F}_{kn}(t_n) = c.$$

Note that for $k = 1$, indeed condition (15) gives the Poisson limiting distribution. In Silverman and Brown (1978) the problem is attacked by looking more generally at U -statistics. They write (page 816) that the Poisson limits of their paper complement the normal limit theorems for U -statistics. It should be expected that the most obvious example, taking as kernel $1(x_1 + \dots + x_k > t)$ leading to (1-) the empirical df of the convolution, should be covered by their results. Unfortunately, it turns out that this is not the case. Assuming (15), as a rule the dependence condition in their paper is not satisfied. It reads as

$$n^{2k-1}P(X_1 + \dots + X_k > t_n, X_1 + \dots + X_{k-1} + X_k^* > t_n) \rightarrow 0, \quad (16)$$

where X_1, \dots, X_k, X_k^* are i.i.d. r.v.'s with df F . Take for instance $k = 2$ and $F = \Phi$. Then Example 3.1 shows that $P(X_1 + \dots + X_k > t_n, X_1 + \dots + X_{k-1} + X_k^* > t_n)$ is of order $\{\overline{F}_k(t_n)\}^{4/3} |\log(\overline{F}_k(t_n))|^{-1/3}$ and hence (15) implies

$$n^{2k-1}P(X_1 + \dots + X_k > t_n, X_1 + \dots + X_{k-1} + X_k^* > t_n) \rightarrow \infty.$$

When taking the exponential distribution, Example 3.2 shows that under (15) again (16) does not hold. In this light it is remarkable that Silverman and Brown's dependence condition is called "a mild condition" in Dabrowski et al. (2002).

One may ask whether not getting a Poisson limit distribution by Theorem A in Silverman and Brown (1978) is due to the method of proof and conditions of that theorem, in the sense that the result still might be true under a weaker condition. In fact, as we will show, this is not the case: the obvious generalization of the Poisson limit for $n\overline{F}_n(t_n)$ to a similar result for $\binom{n}{k}\overline{F}_{kn}(t_n)$ usually does not come true.

The next theorem gives the limiting behavior of $\binom{n}{k}\overline{F}_{kn}$. Firstly, it is seen that the limiting behavior is *not* determined by $\binom{n}{k}\overline{F}_k(t_n)$, as suggested in literature, see (15). Surprisingly, this promising quantity is not the natural parameter to look at. In fact, the suitable parameter is (the even more simple!) quantity $n\overline{F}(t_n/k)$. Note that both can be seen as extensions of $n\overline{F}_n(t_n)$; indeed, both coincide with $n\overline{F}_n(t_n)$ for $k = 1$. Secondly, the theorem gives a picture of the limiting behavior of $\binom{n}{k}\overline{F}_{kn}(t_n)$, similar to the one for $k = 1$, in the sense that according to the limit of $n\overline{F}(t_n/k)$ being 0 or infinity, we get convergence in probability to 0 or infinity. The theorem and the examples following it, show that for t_n 's satisfying (15) the limiting distribution of $\binom{n}{k}\overline{F}_{kn}(t_n)$ degenerates at 0, although its expectation is strictly positive. Moreover, it is exemplified that when the limit of $n\overline{F}(t_n/k)$ is between 0 and infinity we get a non-degenerate limiting distribution, but not a "standard" one and hence, in particular not a Poisson distribution.

Theorem 5 *We have*

$$\binom{n}{k}\overline{F}_{kn}(t_n) \xrightarrow{P} 0 \Leftrightarrow \lim_{n \rightarrow \infty} n\overline{F}\left(\frac{t_n}{k}\right) = 0 \quad (17)$$

and

$$\binom{n}{k}\overline{F}_{kn}(t_n) \xrightarrow{P} \infty \Leftrightarrow \lim_{n \rightarrow \infty} n\overline{F}\left(\frac{t_n}{k}\right) = \infty.$$

Proof. Note that

$$\binom{n}{k}\overline{F}_{kn}(t_n) \xrightarrow{P} 0 \text{ iff } P\left(\binom{n}{k}\overline{F}_{kn}(t_n) = 0\right) \rightarrow 1.$$

We have

$$\begin{aligned} & P\left(\binom{n}{k}\overline{F}_{kn}(t_n) = 0\right) \\ &= P(X_{i_1} + \dots + X_{i_k} \leq t_n \text{ for all } 1 \leq i_1 < \dots < i_k \leq n) \\ &= P(X_{(n-k+1)} + \dots + X_{(n)} \leq t_n) \end{aligned}$$

and hence

$$P\left(X_{(n)} \leq \frac{t_n}{k}\right) \leq P\left(\binom{n}{k}\overline{F}_{kn}(t_n) = 0\right) \leq P\left(X_{(n-k+1)} \leq \frac{t_n}{k}\right). \quad (18)$$

Since

$$P\left(X_{(n)} \leq \frac{t_n}{k}\right) = \left(1 - \overline{F}\left(\frac{t_n}{k}\right)\right)^n$$

and

$$\lim_{n \rightarrow \infty} \left(1 - \overline{F}\left(\frac{t_n}{k}\right)\right)^n = 1 \iff \lim_{n \rightarrow \infty} n\overline{F}\left(\frac{t_n}{k}\right) = 0,$$

it now immediately follows that

$$\lim_{n \rightarrow \infty} n\overline{F}\left(\frac{t_n}{k}\right) = 0$$

implies

$$\lim_{n \rightarrow \infty} P\left(\binom{n}{k}\overline{F}_{kn}(t_n) = 0\right) = 1.$$

On the other hand, if

$$n\overline{F}\left(\frac{t_n}{k}\right) \geq c > 0,$$

then

$$\begin{aligned}
P\left(X_{(n-k+1)} \leq \frac{t_n}{k}\right) &= \sum_{j=0}^k \binom{n}{j} \left\{\overline{F}\left(\frac{t_n}{k}\right)\right\}^j \left\{F\left(\frac{t_n}{k}\right)\right\}^{n-j} \\
&\leq \sum_{j=0}^k \binom{n}{j} \left(\frac{c}{n}\right)^j \left\{1 - \frac{c}{n}\right\}^{n-j} \\
&\leq \sum_{j=0}^k \frac{e^{-c} c^j}{j!} (1 + o(1))
\end{aligned}$$

and hence

$$\lim_{n \rightarrow \infty} P\left(\binom{n}{k} \overline{F}_{kn}(t_n) = 0\right) = 1$$

implies

$$\lim_{n \rightarrow \infty} n \overline{F}\left(\frac{t_n}{k}\right) = 0.$$

This completes the proof of the first part of the theorem.

It is seen from (18) that $\lim_{n \rightarrow \infty} P\left(\binom{n}{k} \overline{F}_{kn}(t_n) = 0\right) = 0$ implies $\lim_{n \rightarrow \infty} P\left(X_{(n)} \leq t_n/k\right) = 0$ and hence $\lim_{n \rightarrow \infty} (1 - \overline{F}(t_n/k))^n = 0$ and thus $\lim_{n \rightarrow \infty} n \overline{F}(t_n/k) = \infty$.

Next assume that $\lim_{n \rightarrow \infty} n \overline{F}(t_n/k) = \infty$. For each integer $A > k$ we have

$$X_{(n-A+1)} \geq t_n/k \Rightarrow \binom{n}{k} \overline{F}_{kn}(t_n) \geq \binom{A}{k}.$$

Because $\lim_{n \rightarrow \infty} n \overline{F}(t_n/k) = \infty$ holds, we get for each integer $A > 0$

$$\lim_{n \rightarrow \infty} \sum_{j=0}^A \binom{n}{j} \left\{\overline{F}\left(\frac{t_n}{k}\right)\right\}^j \left\{F\left(\frac{t_n}{k}\right)\right\}^{n-j} = 0$$

and hence, for each integer $A > k$, we obtain

$$\lim_{n \rightarrow \infty} P\left(X_{(n-A+1)} \geq t_n/k\right) = 1$$

and thus, for each integer $A > k$,

$$\lim_{n \rightarrow \infty} P\left(\binom{n}{k} \overline{F}_{kn}(t_n) \geq \binom{A}{k}\right) = 1,$$

which completes the proof of the second part of the theorem. ■

Remark 4.1 In fact the last statement of Theorem 5 can be extended to

$$\binom{n}{k} \overline{F}_{kn}(t_n) \xrightarrow{P} \infty \Leftrightarrow \lim_{n \rightarrow \infty} P\left(\binom{n}{k} \overline{F}_{kn}(t_n) = 0\right) = 0 \Leftrightarrow \lim_{n \rightarrow \infty} n \overline{F}\left(\frac{t_n}{k}\right) = \infty.$$

This is easily seen from (18) noting that

$$\begin{aligned}
\lim_{n \rightarrow \infty} P\left(X_{(n)} \leq \frac{t_n}{k}\right) = 0 &\Leftrightarrow \lim_{n \rightarrow \infty} n \overline{F}\left(\frac{t_n}{k}\right) = \infty, \\
\lim_{n \rightarrow \infty} P\left(X_{(n-k+1)} \leq \frac{t_n}{k}\right) = 0 &\Leftrightarrow \lim_{n \rightarrow \infty} n \overline{F}\left(\frac{t_n}{k}\right) = \infty.
\end{aligned}$$

Therefore, every limiting distribution of $\binom{n}{k} \overline{F}_{kn}(t_n)$ should have positive probability mass at 0.

Example 4.1 Let $k = 2$ and consider for F the standard normal df Φ . Assume that $\lim_{n \rightarrow \infty} n^2 p_n (\log n)^{-1/2} = 0$. As before, $p_n(1 + \varepsilon_n) = \overline{F}_2(t_n) = \overline{\Phi}(2^{-1/2}t_n)$ with ε_n bounded and hence $\lim_{n \rightarrow \infty} n^2 \overline{\Phi}(2^{-1/2}t_n) (\log n)^{-1/2} = 0$. Since

$$n\overline{\Phi}\left(\frac{t_n}{2}\right) = \left(\frac{2n^2\overline{\Phi}(2^{-1/2}t_n)}{\pi^{1/2}t_n}\right)^{1/2} (1 + o(1)),$$

it is easily seen that $\lim_{n \rightarrow \infty} n^2 \overline{\Phi}(2^{-1/2}t_n) (\log n)^{-1/2} = 0$ implies $\lim_{n \rightarrow \infty} n\overline{\Phi}(t_n/2) = 0$. Application of Theorem 5 yields that $\lim_{n \rightarrow \infty} n^2 p_n (\log n)^{-1/2} = 0$ is a necessary and sufficient condition to get a degenerate limit distribution of $\binom{n}{2} \overline{F}_{2n}(t_n)$ at 0. In particular, for $p_n = n^{-2}$ and $\varepsilon_n = 0$ the limiting distribution of $\binom{n}{2} \overline{F}_{2n}(t_n)$ is not a Poisson distribution, but it degenerates at 0, although (15) holds with $c = \frac{1}{2}$. This shows that the presumed generalization of the convergence to a Poisson distribution for $k \geq 2$ does not come true.

Example 4.2 Let $k = 2$ and consider the exponential distribution with parameter $\lambda = 1$. Assume that $\lim_{n \rightarrow \infty} n^2 p_n (\log n)^{-1} = 0$. Here we have $p_n(1 + \varepsilon_n) = \overline{F}_2(t_n) = (1 + t_n) e^{-t_n}$ and hence, again assuming that ε_n is bounded, we get $\lim_{n \rightarrow \infty} n^2 t_n e^{-t_n} (\log n)^{-1} = 0$. Since

$$\begin{aligned} n\overline{F}\left(\frac{t_n}{2}\right) &= n \exp\left(-\frac{t_n}{2}\right) \\ &= \left(\frac{n^2 t_n \exp(-t_n)}{\log n}\right)^{1/2} \left(\frac{\log n}{t_n}\right)^{1/2}, \end{aligned}$$

it easily follows that $\lim_{n \rightarrow \infty} n^2 p_n (\log n)^{-1} = 0$ implies $\lim_{n \rightarrow \infty} n\overline{F}(t_n/2) = 0$. Application of Theorem 5 yields that $\lim_{n \rightarrow \infty} n^2 p_n (\log n)^{-1} = 0$ is a necessary and sufficient condition to get a degenerate limit distribution of $\binom{n}{2} \overline{F}_{2n}(t_n)$ at 0. In particular, for $p_n = n^{-2}$ and $\varepsilon_n = 0$ the limiting distribution of $\binom{n}{2} \overline{F}_{2n}(t_n)$ is not a Poisson distribution, but it degenerates at 0, although (15) holds with $c = \frac{1}{2}$. Again this shows that the presumed generalization of the convergence to a Poisson distribution for $k \geq 2$ does not come true.

Remark 4.2 When showing that convergence in probability does not imply convergence of moments, many textbooks give examples of the type: take a r.v. with probability $1 - n^{-1}$ at 0 and probability n^{-1} at n . Such examples may look artificial, but Examples 4.1 and 4.2 are "real" examples with the same flavor: probability mass tending to 1 at the point 0 and expectation equal to $\frac{1}{2}$. Both for the normal and for the exponential distribution p_n can even be chosen such that $\binom{n}{2} \overline{F}_{2n}(t_n)$ converges in probability to 0, while its expectation converges to ∞ .

Remark 4.3 Let u_n satisfy

$$\lim_{n \rightarrow \infty} n\overline{F}(u_n) = 0, \quad (19)$$

implying that $\lim_{n \rightarrow \infty} P(\max_{1 \leq i \leq n} X_i \leq u_n) = 1$. As noted in Remark 3.4 we may restrict attention to $X \leq u_n$. By an obvious adaptation of the proof, convergence of $\binom{n}{k} \overline{F}_{kn}(t_n)$ to 0 in probability continuous to hold when condition (17) is replaced by

$$\lim_{n \rightarrow \infty} \min \left\{ n\overline{F}\left(\frac{t_n - ju_n}{k - j}\right) : j = 0, \dots, k - 1 \right\} = 0. \quad (20)$$

Condition (20) seems to be a weaker condition than condition (17), since of course

$$\min \left\{ n\overline{F}\left(\frac{t_n - ju_n}{k - j}\right) : j = 0, \dots, k - 1 \right\} \leq n\overline{F}\left(\frac{t_n}{k}\right).$$

However, assuming (19), condition (20) and condition (17) are in fact equivalent, since for any $j = 0, \dots, k - 1$

$$\frac{t_n}{k} \geq \min \left\{ u_n, \frac{t_n - ju_n}{k - j} \right\}.$$

Therefore, restricting to $X \leq u_n$ gives no improvement. Indeed, no improvement could be expected, because (17) is a necessary and sufficient condition.

The following example illustrates the behavior of $\binom{n}{k} \overline{F}_{kn}(t_n)$ when $\lim_{n \rightarrow \infty} n \overline{F}(t_n/k) = c > 0$.

Example 4.3 Let $k = 2$ and consider the exponential distribution with parameter $\lambda = 1$. Assume that $\lim_{n \rightarrow \infty} n \overline{F}(t_n/2) = c > 0$. We have, writing f for the density,

$$\begin{aligned} & P \left(\binom{n}{2} \overline{F}_{2n}(t_n) = 0 \right) \\ &= P(X_{(n-1)} + X_{(n)} \leq t_n) \\ &= P \left(X_{(n-1)} \leq \frac{t_n}{2} \right) - P \left(X_{(n-1)} \leq \frac{t_n}{2}, X_{(n-1)} + X_{(n)} > t_n \right) \\ &= \left\{ F \left(\frac{t_n}{2} \right) \right\}^n + \binom{n}{1} \left\{ F \left(\frac{t_n}{2} \right) \right\}^{n-1} \overline{F} \left(\frac{t_n}{2} \right) - \int_{-\infty}^{t_n/2} n(n-1) f(x) \{F(x)\}^{n-2} \overline{F}(t_n - x) dx. \end{aligned} \quad (21)$$

Using $\lim_{n \rightarrow \infty} n \overline{F}(t_n/2) = c > 0$ we obtain

$$\begin{aligned} \lim_{n \rightarrow \infty} \left\{ F \left(\frac{t_n}{2} \right) \right\}^n &= \lim_{n \rightarrow \infty} \left\{ 1 - \overline{F} \left(\frac{t_n}{2} \right) \right\}^n = e^{-c}, \\ \lim_{n \rightarrow \infty} \binom{n}{1} \left\{ F \left(\frac{t_n}{2} \right) \right\}^{n-1} \overline{F} \left(\frac{t_n}{2} \right) &= \lim_{n \rightarrow \infty} n \overline{F}(t_n/2) \left\{ 1 - \overline{F} \left(\frac{t_n}{2} \right) \right\}^{n-1} = ce^{-c}. \end{aligned} \quad (22)$$

Next we use the particular form of the distribution here, that is of the exponential distribution, to find the limit of the integral in (21). We get

$$\begin{aligned} & \int_{-\infty}^{t_n/2} n(n-1) f(x) \{F(x)\}^{n-2} \overline{F}(t_n - x) dx \\ &= \int_0^{t_n/2} n(n-1) e^{-t_n} (1 - e^{-x})^{n-2} dx. \end{aligned} \quad (23)$$

Substitution of $y = ne^{-x}$ gives

$$\begin{aligned} & \int_0^{t_n/2} (1 - e^{-x})^{n-2} dx \\ &= \int_{n \exp(-t_n/2)}^n y^{-1} \left(1 - \frac{y}{n} \right)^{n-2} dy. \end{aligned}$$

For all $n \geq n_1 \geq 4$ we have $ne^{-t_n/2} \geq c/2$ and $(1 - \frac{y}{n})^{n-2} \leq (1 - \frac{y}{n})^{n/2}$. Hence, for $n \geq n_1$ we get

$$\begin{aligned} & y^{-1} \left(1 - \frac{y}{n} \right)^{n-2} \mathbf{1} \left(ne^{-t_n/2} < y < n \right) \\ & \leq y^{-1} \left(1 - \frac{y}{n} \right)^{n/2} \mathbf{1} \left(c/2 < y < n \right) \leq y^{-1} e^{-y/2} \mathbf{1} \left(c/2 < y < \infty \right), \end{aligned}$$

which is an integrable function. Application of the dominated convergence theorem yields

$$\begin{aligned} & \lim_{n \rightarrow \infty} \int_{n \exp(-t_n/2)}^n y^{-1} \left(1 - \frac{y}{n}\right)^{n-2} dy \\ &= \int_c^\infty \frac{e^{-y}}{y} dy \end{aligned}$$

and therefore

$$\begin{aligned} & \lim_{n \rightarrow \infty} \int_0^{t_n/2} n(n-1)e^{-t_n} (1 - e^{-x})^{n-2} dx \\ &= c^2 \int_c^\infty \frac{e^{-y}}{y} dy. \end{aligned} \tag{24}$$

Combination of (21) – (24) gives

$$\lim_{n \rightarrow \infty} P \left(\binom{n}{2} \overline{F}_{2n}(t_n) = 0 \right) = e^{-c} (1 + c) - c^2 \int_c^\infty \frac{e^{-y}}{y} dy.$$

Because

$$c^2 \int_c^\infty \frac{e^{-y}}{y} dy < c \int_c^\infty e^{-y} dy = ce^{-c},$$

we obtain

$$\lim_{n \rightarrow \infty} P \left(\binom{n}{2} \overline{F}_{2n}(t_n) = 0 \right) > e^{-c}$$

and therefore the limiting distribution is not the Poisson distribution with parameter c .

Furthermore, we get

$$\begin{aligned} & P \left(\binom{n}{2} \overline{F}_{2n}(t_n) = 1 \right) \\ &= P(X_{(n-2)} + X_{(n)} < t_n, X_{(n-1)} + X_{(n)} > t_n) \\ &= P(X_{(n-2)} < t_n/2, X_{(n-1)} > t_n/2) - P(X_{(n-2)} < t_n/2, X_{(n-2)} + X_{(n)} > t_n) \\ &+ P(X_{(n-1)} < t_n/2, X_{(n-1)} + X_{(n)} > t_n) \end{aligned}$$

and in a similar way as above we arrive at

$$\begin{aligned} & \lim_{n \rightarrow \infty} P \left(\binom{n}{2} \overline{F}_{2n}(t_n) = 1 \right) \\ &= \frac{c^2}{2} e^{-c} (c - 1) + \frac{c^2}{2} (2 - c^2) \int_c^\infty \frac{e^{-y}}{y} dy. \end{aligned}$$

Writing

$$f_0(c) = \lim_{n \rightarrow \infty} P \left(\binom{n}{2} \overline{F}_{2n}(t_n) = 0 \right), f_1(c) = \lim_{n \rightarrow \infty} P \left(\binom{n}{2} \overline{F}_{2n}(t_n) = 1 \right)$$

it turns out that for all $c > 0$

$$f_1(c) < f_0(c) \{-\log(f_0(c))\}$$

and hence, whatever the parameterization and whatever c , the limiting distribution is not a Poisson distribution.

Summarizing the results of this section, we see that asymptotic degeneration of $\binom{n}{k} \overline{F}_{kn}(t_n)$ at 0 appears iff $\lim_{n \rightarrow \infty} n \overline{F}(t_n/k) = 0$, that a non-degenerate limit distribution may come up

when $\lim_{n \rightarrow \infty} n\bar{F}(t_n/k) = c > 0$, and finally we see that $\lim_{n \rightarrow \infty} n\bar{F}(t_n/k) = \infty$ is equivalent to convergence in probability to ∞ . In the latter case a standardization is conceivable leading to a non-degenerate limit distribution. Indeed, under suitable conditions asymptotic normality can be achieved as is shown in Section 3.

The Examples 3.3 and 3.4 show that not the whole range given by $\lim_{n \rightarrow \infty} n\bar{F}(t_n/k) = \infty$ is covered by the results on asymptotic normality. Let $k = 2$. For the standard normal distribution $\lim_{n \rightarrow \infty} np_n (\log n)^{1/2} = \infty$ ensures asymptotic normality, while $\lim_{n \rightarrow \infty} n\bar{\Phi}(t_n/2) = \infty$ corresponds to $\lim_{n \rightarrow \infty} n^2 p_n (\log n)^{-1/2} = \infty$. In case of the exponential distribution asymptotic normality is guaranteed if $\lim_{n \rightarrow \infty} np_n (\log n)^{-1} = \infty$, while $\lim_{n \rightarrow \infty} n\bar{F}(t_n/2) = \infty$ is equivalent to $\lim_{n \rightarrow \infty} n^2 p_n (\log n)^{-1} = \infty$. Whether another standardization leads to a (standard) non-degenerate limit distribution for the remaining p_n 's is unknown. In view of the sharpness of the tools used in Section 3 it is doubtful that (with the natural standardization used in Section 3) asymptotic normality can be extended to a larger range of p_n 's.

5 Exact inequalities

Exact inequalities may be used to get the exceedance probability at most equal to α , that is

$$P(P_n(q_n) > p_n(1 + \varepsilon_n)) \leq \alpha.$$

In view of Lemma 1 we can equivalently require

$$P\left(\bar{F}_{kn}\left(\bar{F}_k^{-1}(p_n(1 + \varepsilon_n))\right) \leq \tilde{q}_n\right) \leq \alpha, \text{ with } \tilde{q}_n = \frac{\left[\binom{n}{k} q_n\right]}{\binom{n}{k}}.$$

The following theorem is based on the Markov and Chebyshev inequalities. Write $x^+ = \max(x, 0)$.

Theorem 6 *Let*

$$U_n = p_n(1 + \varepsilon_n) - \bar{F}_{kn}\left(\bar{F}_k^{-1}(p_n(1 + \varepsilon_n))\right).$$

For any $r > 0$ and

$$\tilde{q}_n \leq p_n(1 + \varepsilon_n) - \left[\frac{1}{\alpha} E(\{U_n^+\}^r)\right]^{1/r} \quad (25)$$

it holds that

$$P\left(\bar{F}_{kn}\left(\bar{F}_k^{-1}(p_n(1 + \varepsilon_n))\right) \leq \tilde{q}_n\right) \leq \alpha.$$

In particular, if

$$\tilde{q}_n \leq p_n(1 + \varepsilon_n) - \sqrt{\frac{1}{\alpha} \text{var}\left(\bar{F}_{kn}\left(\bar{F}_k^{-1}(p_n(1 + \varepsilon_n))\right)\right)}, \quad (26)$$

then

$$P\left(\bar{F}_{kn}\left(\bar{F}_k^{-1}(p_n(1 + \varepsilon_n))\right) \leq \tilde{q}_n\right) \leq \alpha. \quad (27)$$

Proof. For any r.v. Z we have

$$P(Z \geq t) \leq \frac{E(\{Z^+\}^r)}{t^r}$$

for any $r, t > 0$. Application of this inequality with $Z = U_n$ and $t = p_n(1 + \varepsilon_n) - \tilde{q}_n$ yields

$$\begin{aligned} & P\left(\bar{F}_{kn}\left(\bar{F}_k^{-1}(p_n(1 + \varepsilon_n))\right) \leq \tilde{q}_n\right) \\ &= P(U_n \geq p_n(1 + \varepsilon_n) - \tilde{q}_n) \\ &\leq \frac{E(\{U_n^+\}^r)}{(p_n(1 + \varepsilon_n) - \tilde{q}_n)^r} \leq \alpha, \end{aligned}$$

where the last inequality follows from condition (25).

By taking $r = 2$ and using $E\left(\{U_n^+\}^2\right) \leq E(U_n^2) = \text{var}(U_n)$, it immediately follows that (26) implies (25) with $r = 2$ and hence (27) holds. ■

Let $X_1, \dots, X_k, X_1^*, \dots, X_k^*$ be i.i.d. r.v.'s with df F . For any $1 \leq j \leq k$ define

$$h_{jk}(t) = P(X_1 + \dots + X_j + X_{j+1} + \dots + X_k > t, X_1 + \dots + X_j + X_{j+1}^* + \dots + X_k^* > t)$$

and

$$\sigma_{kn}^2(t) = \frac{1}{\binom{n}{k}} \sum_{j=1}^k \binom{k}{j} \binom{n-k}{k-j} \left[h_{jk}(t) - \{\bar{F}_k(t)\}^2 \right].$$

In view of (3) we have

$$h_{1k}(t) - \{\bar{F}_k(t)\}^2 = \{s_n(t)\}^2.$$

Moreover,

$$h_{kk}(t) = \bar{F}_k(t) - \{\bar{F}_k(t)\}^2$$

and hence

$$\sigma_{2n}^2(t) = \frac{4(n-2)}{n(n-1)} \{s_n(t)\}^2 + \frac{2}{n(n-1)} \bar{F}_2(t) \{1 - \bar{F}_2(t)\}.$$

Corollary 7 Let $t_n = \bar{F}_k^{-1}(p_n(1 + \varepsilon_n))$. If

$$\tilde{q}_n \leq p_n(1 + \varepsilon_n) - \sigma_{kn}(t_n) \sqrt{\frac{1}{\alpha}}, \quad (28)$$

then

$$P\left(\bar{F}_{kn}\left(\bar{F}_k^{-1}(p_n(1 + \varepsilon_n))\right) \leq \tilde{q}_n\right) \leq \alpha.$$

Proof. By Lemma A on page 183 of Serfling (1980) and some direct calculation it is seen that

$$\text{var}\left(\bar{F}_{kn}\left(\bar{F}_k^{-1}(p_n(1 + \varepsilon_n))\right)\right) = \sigma_{kn}^2(t_n).$$

The result now immediately follows from the last part of Theorem 6. ■

In the following examples we compare the correction terms obtained from Corollary 7 with those from Theorem 4.

Example 5.1 Let $k = 2$ and consider for F the standard normal df Φ . Writing $\eta_n = p_n(1 + \varepsilon_n) = \bar{F}_2(t_n) = \bar{\Phi}(2^{-1/2}t_n)$ with ε_n bounded, it is shown in Example 3.1 that

$$\{s_n(t_n)\}^2 = 3^{3/2} 2^{-5/3} \pi^{-1/3} \eta_n^{4/3} |\log \eta_n|^{-1/3} (1 + o(1)),$$

implying

$$\begin{aligned} \sigma_{2n}^2(t_n) &= \frac{4(n-2)}{n(n-1)} \{s_n(t_n)\}^2 + \frac{2}{n(n-1)} \eta_n \{1 - \eta_n\} \\ &= 3^{3/2} 2^{1/3} \pi^{-1/3} n^{-1} \eta_n^{4/3} |\log \eta_n|^{-1/3} (1 + o(1)), \end{aligned}$$

provided that $\lim_{n \rightarrow \infty} n^3 p_n |\log p_n|^{-1} = \infty$. If $\lim_{n \rightarrow \infty} n^2 p_n (\log n)^{-1/2} = 0$ we get $\binom{n}{2} \bar{F}_{2n}(t_n) \xrightarrow{P} 0$ (see Example 4.1) and therefore we assume further on that $\liminf_{n \rightarrow \infty} n^2 p_n (\log n)^{-1/2} > 0$ (implying $\lim_{n \rightarrow \infty} n^3 p_n |\log p_n|^{-1} = \infty$). Inequality (28) now reads as

$$q_n \leq p_n(1 + \varepsilon_n) - 3^{3/4} 2^{1/6} \pi^{-1/6} n^{-1/2} \eta_n^{2/3} |\log \eta_n|^{-1/6} \sqrt{\frac{1}{\alpha}} (1 + o(1)). \quad (29)$$

(Note that we may replace \tilde{q}_n by q_n here, since $n^{-2} = o(n^{-1/2}\eta_n^{2/3}|\log \eta_n|^{-1/6})$.) Comparison with Example 3.5, in particular with (13), shows that the correction term is of the same order, but that the constant differs: $\overline{\Phi}^{-1}(\alpha)$ is replaced by $\alpha^{-1/2}$. This is obviously no surprise in view of the way both equations are derived, taking into account that $4n^{-1}\{s_n(t_n)\}^2$ is the dominating term in $\sigma_{2n}^2(t_n)$. However, here there is no restriction on p_n , apart from the obvious condition $\liminf_{n \rightarrow \infty} n^2 p_n (\log n)^{-1/2} > 0$.

Example 5.2 Let $k = 2$ and consider the exponential distribution with parameter $\lambda = 1$. Writing $\eta_n = p_n(1 + \varepsilon_n) = \overline{F}_2(t_n) = (1 + t_n)e^{-t_n}$ with ε_n bounded, it is shown in Example 3.2 that

$$\begin{aligned}\{s_n(t_n)\}^2 &= 2e^{-t_n} - e^{-2t_n} - (1 + t_n)^2 e^{-2t_n} \\ &= 2e^{-t_n}(1 + o(1)) \\ &= 2\eta_n |\log \eta_n|^{-1}(1 + o(1))\end{aligned}$$

and hence

$$\begin{aligned}\sigma_{2n}^2(t_n) &= \frac{4(n-2)}{n(n-1)} \{s_n(t_n)\}^2 + \frac{2}{n(n-1)} \eta_n \{1 - \eta_n\} \\ &= 8n^{-1} \eta_n |\log \eta_n|^{-1}(1 + o(1)),\end{aligned}$$

provided that $\lim_{n \rightarrow \infty} n |\log p_n|^{-1} = \infty$. If $\lim_{n \rightarrow \infty} n^2 p_n (\log n)^{-1} = 0$ we get $\binom{n}{2} \overline{F}_{2n}(t_n) \xrightarrow{P} 0$ (see Example 4.2) and therefore we assume further on that $\liminf_{n \rightarrow \infty} n^2 p_n (\log n)^{-1} > 0$ (implying $\lim_{n \rightarrow \infty} n |\log p_n|^{-1} = \infty$). Inequality (28) now reads as

$$q_n \leq p_n(1 + \varepsilon_n) - 2^{3/2} n^{-1/2} \eta_n^{1/2} |\log \eta_n|^{-1/2} \sqrt{\frac{1}{\alpha}}(1 + o(1)). \quad (30)$$

(Note that we may replace \tilde{q}_n by q_n here, since $n^{-2} = o(n^{-1/2}\eta_n^{1/2}|\log \eta_n|^{-1/2})$.) Comparison with Example 3.6, in particular with (14), shows that the correction term is of the same order, but again of course the constant $\overline{\Phi}^{-1}(\alpha)$ is replaced by $\alpha^{-1/2}$. Here there is no restriction on p_n , apart from the obvious condition $\liminf_{n \rightarrow \infty} n^2 p_n (\log n)^{-1} > 0$.

Remark 5.1 Similarly as in Remark 3.2 some improvement can be made here by replacing F by the df of X given $X \leq u_n$, where u_n satisfies

$$\lim_{n \rightarrow \infty} n \overline{F}(u_n) = 0,$$

and thus

$$\lim_{n \rightarrow \infty} P\left(\max_{1 \leq i \leq n} X_i \leq u_n\right) = 1.$$

Note that in applying Chebyshev's inequality in this situation, the starting point is no longer $p_n(1 + \varepsilon_n) = \overline{F}_k\left(\overline{F}_k^{-1}(p_n(1 + \varepsilon_n))\right)$, but (the smaller) $\overline{H}_k\left(\overline{F}_k^{-1}(p_n(1 + \varepsilon_n))\right)$, where H_k is the df of the convolution of k r.v.'s each with as df the df of X given $X \leq u_n$.

Next we consider some Bernstein type inequalities. The most well-known is due to Hoeffding (1963), see also Serfling (1980) page 201, and reads in our case as

$$\begin{aligned}P(\sqrt{n} \{\overline{F}_{kn}(t) - \overline{F}_k(t)\} \geq x) \\ \leq \exp \left\{ - \frac{x^2}{2k [\overline{F}_k(t) (1 - \overline{F}_k(t)) + 3^{-1} (1 - \overline{F}_k(t)) x n^{-1/2}]} \right\}.\end{aligned}$$

Arcones (1995) page 240 rightly remarks that the "variance-term" $\overline{F}_k(t) (1 - \overline{F}_k(t))$ appearing here does not coincide with the variance of the limiting normal distribution, see also Corollary 3. Indeed, in our situation $\{s_n(t)\}^2$ is usually of a lower order than $\overline{F}_k(t)$. Therefore, Hoeffding's inequality is not appropriate for our application. Fortunately, Theorem 2 of Arcones (1995) gives a Bernstein-type inequality with $\overline{F}_k(t) (1 - \overline{F}_k(t))$ replaced by $\{s_n(t)\}^2$.

Theorem 8 For all $x > 0$ and $t \in \mathbb{R}$, we have

$$\begin{aligned} & P\left(\frac{\sqrt{n}}{k}T_n(t) \leq -x\right) \\ & \leq 2 \exp\left\{-\frac{x^2}{2 + \{2^{k+3}k^k(n-1)^{-1/2} + (2/3)n^{-1/2}\}x\{s_n(t)\}^{-1}}\right\}. \end{aligned} \quad (31)$$

Proof. By definition of T_n we have

$$P\left(\frac{\sqrt{n}}{k}T_n(t) \leq -x\right) = P\left(\sqrt{n}\{\overline{F}_{kn}(t) - \overline{F}_k(t)\} \leq -xks_n(t)\right).$$

Application of Theorem 2 in Arcones (1995) (with the factor 4 replaced by 2, since we have a one-sided inequality) gives the result, where we note that in (2.10) in Arcones (1995) one should read $2^{m+3}m^{m+1}$ instead of $2^{m+2}m^m$ and m instead of m^{-1} . ■

Corollary 9 Let $t_n = \overline{F}_k^{-1}(p_n(1 + \varepsilon_n))$. If

$$\lim_{n \rightarrow \infty} n\{s_n(t_n)\}^2 = \infty, \quad (32)$$

then, for all $x > 0$,

$$P\left(\frac{\sqrt{n}}{k}T_n(t_n) \leq -x\right) \leq 2 \exp\left(-\frac{1}{2}x^2\right)(1 + o(1)) \quad (33)$$

as $n \rightarrow \infty$. If moreover

$$\tilde{q}_n \leq p_n(1 + \varepsilon_n) - s_n\left(\overline{F}_k^{-1}(p_n(1 + \varepsilon_n))\right)kn^{-1/2}\sqrt{2|\ln(\alpha/2)|}, \quad (34)$$

then

$$\limsup_{n \rightarrow \infty} P\left(\overline{F}_{kn}\left(\overline{F}_k^{-1}(p_n(1 + \varepsilon_n))\right) \leq \tilde{q}_n\right) \leq \alpha. \quad (35)$$

Proof. The first result easily follows from Theorem 8, since (32) implies

$$\left\{2^{k+3}k^k(n-1)^{-1/2} + (2/3)n^{-1/2}\right\}x\{s_n(t_n)\}^{-1} = o(1)$$

as $n \rightarrow \infty$. Writing $t_n = \overline{F}_k^{-1}(p_n(1 + \varepsilon_n))$, we have

$$\begin{aligned} & P\left(\overline{F}_{kn}\left(\overline{F}_k^{-1}(p_n(1 + \varepsilon_n))\right) \leq \tilde{q}_n\right) \\ & \leq P\left(\frac{\sqrt{n}}{k}T_n(t_n) \leq -\sqrt{2|\ln(\alpha/2)|}\right) \end{aligned}$$

and (35) follows from (33). ■

To shed some light on the benefits of the Bernstein inequality (Corollary 9) compared to Chebyshev's inequality (Corollary 7) and the asymptotic normality (Theorem 4) we consider the examples of the standard normal and exponential distribution. Condition (32) may be rather

restrictive, see also Examples 5.3 and 5.4. On the other hand, if $\lim_{n \rightarrow \infty} n \{s_n(t_n)\}^2 = 0$, clearly the inequality is not very useful, because in that case we should take $x = x_n$ tending to infinity (at the rate $n^{-1/2} \{s_n(t)\}^{-1}$).

Example 5.3 Let $k = 2$ and consider for F the standard normal df Φ . Let ε_n be bounded. Assume that $\lim_{n \rightarrow \infty} n \{s_n(t_n)\}^2 = \infty$, or (see Example 3.1), equivalently, $\lim_{n \rightarrow \infty} np_n^{4/3} |\log p_n|^{-1/3} = \infty$. For those p_n 's Chebyshev's inequality gives as correction term, see (29),

$$-2^{1/6} 3^{3/4} \pi^{-1/6} n^{-1/2} \eta_n^{2/3} |\log \eta_n|^{-1/6} \sqrt{\frac{1}{\alpha}}.$$

Asymptotic normality yields as correction term, see (13),

$$-2^{1/6} 3^{3/4} \pi^{-1/6} n^{-1/2} \eta_n^{2/3} |\log \eta_n|^{-1/6} \bar{\Phi}^{-1}(\alpha).$$

Bernstein's inequality implies the correction term, see (34),

$$-2^{1/6} 3^{3/4} \pi^{-1/6} n^{-1/2} \eta_n^{2/3} |\log \eta_n|^{-1/6} \sqrt{2 |\ln(\alpha/2)|}.$$

We have for $0 < \alpha < 0.23$

$$\sqrt{\frac{1}{\alpha}} \geq \sqrt{2 |\ln(\alpha/2)|} \geq \bar{\Phi}^{-1}(\alpha)$$

and hence Bernstein's inequality gives some improvement w.r.t. Chebyshev's inequality in this (restricted) range of p_n 's.

Example 5.4 Let $k = 2$ and consider the exponential distribution with parameter $\lambda = 1$. Let ε_n be bounded. Assume that $\lim_{n \rightarrow \infty} n \{s_n(t_n)\}^2 = \infty$, or (see Example 3.2), equivalently, $\lim_{n \rightarrow \infty} np_n |\log p_n|^{-1} = \infty$. For those p_n 's Chebyshev's inequality, asymptotic normality and Bernstein's inequality give the same correction terms apart from the factors $\sqrt{\alpha^{-1}}$, $\sqrt{2 |\ln(\alpha/2)|}$ and $\bar{\Phi}^{-1}(\alpha)$, see (30), (14) and (34). Hence, similarly as in Example 5.3, Bernstein's inequality gives some improvement w.r.t. Chebyshev's inequality in this range of p_n 's.

Appendix Proof of (7)

Let (X, Y) have a standard bivariate normal distribution with correlation coefficient ρ , $-1 < \rho < 1$. Then $Y|X = x$ has a normal distribution with expectation ρx and variance $1 - \rho^2$. Hence,

$$\begin{aligned} P(Y > t | X > t) &= \{\bar{\Phi}(t)\}^{-1} \int_t^\infty \bar{\Phi}\left(\frac{t - \rho x}{\sqrt{1 - \rho^2}}\right) \varphi(x) dx \\ &= \{\bar{\Phi}(t)\}^{-1} \int_t^\infty 1(t - \rho x \geq \sqrt{t}) \bar{\Phi}\left(\frac{t - \rho x}{\sqrt{1 - \rho^2}}\right) \varphi(x) dx \\ &\quad + \{\bar{\Phi}(t)\}^{-1} \int_t^\infty 1(t - \rho x < \sqrt{t}) \bar{\Phi}\left(\frac{t - \rho x}{\sqrt{1 - \rho^2}}\right) \varphi(x) dx. \end{aligned}$$

We use the following well-known properties

$$\begin{aligned} \lim_{x \rightarrow \infty} \sup_{y \geq x} \left| \frac{y \bar{\Phi}(y)}{\varphi(y)} - 1 \right| &= 0, \\ \lim_{x \rightarrow \infty} \frac{x^j \bar{\Phi}(x + c)}{\bar{\Phi}(x)} &= 0 \end{aligned}$$

for any $j, c > 0$. Firstly, we show that

$$\{\bar{\Phi}(t)\}^{-1} \int_t^\infty 1(t - \rho x < \sqrt{t}) \bar{\Phi}\left(\frac{t - \rho x}{\sqrt{1 - \rho^2}}\right) \varphi(x) dx = o\left(\bar{\Phi}\left(t\sqrt{\frac{1 - \rho}{1 + \rho}}\right)\right).$$

Obviously, for $\rho \leq 0$

$$\int_t^\infty 1(t - \rho x < \sqrt{t}) \bar{\Phi}\left(\frac{t - \rho x}{\sqrt{1 - \rho^2}}\right) \varphi(x) dx = 0.$$

For $\rho > 0$ we have

$$\begin{aligned} & \{\bar{\Phi}(t)\}^{-1} \int_t^\infty 1(t - \rho x < \sqrt{t}) \bar{\Phi}\left(\frac{t - \rho x}{\sqrt{1 - \rho^2}}\right) \varphi(x) dx \\ & \leq \{\bar{\Phi}(t)\}^{-1} \int_{(t - \sqrt{t})\rho^{-1}}^\infty \varphi(x) dx = \{\bar{\Phi}(t)\}^{-1} \bar{\Phi}\left(\frac{t - \sqrt{t}}{\rho}\right) = o\left(\bar{\Phi}\left(t\sqrt{\frac{1 - \rho}{1 + \rho}}\right)\right). \end{aligned}$$

Next we consider

$$\begin{aligned} & \{\bar{\Phi}(t)\}^{-1} \int_t^\infty 1(t - \rho x \geq \sqrt{t}) \bar{\Phi}\left(\frac{t - \rho x}{\sqrt{1 - \rho^2}}\right) \varphi(x) dx \\ & = \frac{t}{\varphi(t)} \int_t^\infty 1(t - \rho x \geq \sqrt{t}) \frac{\sqrt{1 - \rho^2}}{t - \rho x} \varphi\left(\frac{t - \rho x}{\sqrt{1 - \rho^2}}\right) \varphi(x) dx (1 + o(1)) \\ & = t \int_t^\infty 1(t - \rho x \geq \sqrt{t}) \frac{\sqrt{1 - \rho^2}}{t - \rho x} \varphi\left(\frac{x - \rho t}{\sqrt{1 - \rho^2}}\right) dx (1 + o(1)) \\ & = (1 + \rho) \int_t^\infty \frac{1}{\sqrt{1 - \rho^2}} \varphi\left(\frac{x - \rho t}{\sqrt{1 - \rho^2}}\right) dx (1 + o(1)) \\ & + \int_t^{t+1} \left\{ \frac{t(1 - \rho^2)}{t - \rho x} 1(t - \rho x \geq \sqrt{t}) - (1 + \rho) \right\} \frac{1}{\sqrt{1 - \rho^2}} \varphi\left(\frac{x - \rho t}{\sqrt{1 - \rho^2}}\right) dx (1 + o(1)) \\ & + \int_{t+1}^\infty \left\{ \frac{t(1 - \rho^2)}{t - \rho x} 1(t - \rho x \geq \sqrt{t}) - (1 + \rho) \right\} \frac{1}{\sqrt{1 - \rho^2}} \varphi\left(\frac{x - \rho t}{\sqrt{1 - \rho^2}}\right) dx (1 + o(1)). \end{aligned}$$

We have, for $-1 < \rho < 1$,

$$\begin{aligned} & (1 + \rho) \int_t^\infty \frac{1}{\sqrt{1 - \rho^2}} \varphi\left(\frac{x - \rho t}{\sqrt{1 - \rho^2}}\right) dx (1 + o(1)) \\ & = (1 + \rho) \bar{\Phi}\left(\frac{t - \rho t}{\sqrt{1 - \rho^2}}\right) (1 + o(1)) \\ & = (1 + \rho) \bar{\Phi}\left(t\sqrt{\frac{1 - \rho}{1 + \rho}}\right) (1 + o(1)) \end{aligned}$$

and

$$\begin{aligned} & \int_t^{t+1} \left\{ \frac{t(1 - \rho^2)}{t - \rho x} 1(t - \rho x \geq \sqrt{t}) - (1 + \rho) \right\} \frac{1}{\sqrt{1 - \rho^2}} \varphi\left(\frac{x - \rho t}{\sqrt{1 - \rho^2}}\right) dx (1 + o(1)) \\ & = o\left(\int_t^{t+1} \frac{1}{\sqrt{1 - \rho^2}} \varphi\left(\frac{x - \rho t}{\sqrt{1 - \rho^2}}\right) dx\right) = o\left(\bar{\Phi}\left(t\sqrt{\frac{1 - \rho}{1 + \rho}}\right)\right), \end{aligned}$$

while

$$\begin{aligned}
& \left| \int_{t+1}^{\infty} \left\{ \frac{t(1-\rho^2)}{t-\rho x} 1(t-\rho x \geq \sqrt{t}) - (1+\rho) \right\} \frac{1}{\sqrt{1-\rho^2}} \varphi\left(\frac{x-\rho t}{\sqrt{1-\rho^2}}\right) dx (1+o(1)) \right| \\
& \leq \int_{t+1}^{\infty} \left\{ \frac{t(1-\rho^2)}{\sqrt{t}} 1(t-\rho x \geq \sqrt{t}) + |1+\rho| \right\} \frac{1}{\sqrt{1-\rho^2}} \varphi\left(\frac{x-\rho t}{\sqrt{1-\rho^2}}\right) dx (1+o(1)) \\
& \leq \left\{ \sqrt{t} + |1+\rho| \right\} \bar{\Phi}\left(\frac{t-\rho t}{\sqrt{1-\rho^2}} + \frac{1}{\sqrt{1-\rho^2}}\right) (1+o(1)) \\
& = o\left(\bar{\Phi}\left(t\sqrt{\frac{1-\rho}{1+\rho}}\right)\right).
\end{aligned}$$

This completes the proof of (7).

References

- Albers, W. and Kallenberg, W.C.M. (2004a). Are estimated control charts in control? *Statistics* **38** 67-79.
- Albers, W. and Kallenberg, W.C.M. (2004b). Empirical nonparametric control charts: estimation effects and corrections. *J. Appl. Stat.* **31** 345-360.
- Albers, W., Kallenberg, W.C.M. and Nurdiati, S. (2002). Exceedance probabilities for parametric control charts. Technical Report 1650, University of Twente.
- Albers, W., Kallenberg, W.C.M. and Nurdiati, S. (2004). Parametric control charts. *J. Statist. Plann. Inference* **124** 159-184.
- Arcones, M.A. (1995). A Bernstein-type inequality for U -statistics and U -processes. *Statist. Probab. Lett.* **22** 239-247.
- Arcones, M.A. (1996). The Bahadur-Kiefer representation for U -quantiles. *Ann. Stat.* **24** 1400-1422.
- Bentkus, V. Götze, F. and Zitikis, R. (1994). Lower estimates of the convergence rate for U -statistics. *Ann. Probab.* **22** 1707-1714.
- Choudhury, J. and Serfling, R.J. (1988). Generalized order statistics, Bahadur representations and sequential nonparametric fixed-width confidence intervals. *J. Statist. Plann. Inference* **19** 269-282.
- Dabrowski, A.R. Dehling, H.G., Mikosch, T. and Sharipov, O. (2002) Poisson limits for U -statistics *Stochastic Process. Appl.* **99** 137-157.
- Embrechts, P., McNeil, A.J. and Straumann, D. (2002) Correlation and dependence in risk management: properties and pitfalls. In: *Risk management: value at risk and beyond*. (M.A.H. Dempster ed.) 176-223. Cambridge Univ. Press.
- Friedrich, K.O. (1989). A Berry-Esseen bound for functions of independent random variables. *Ann. Stat.* **17** 170-183.
- Hoeffding, W. (1963). Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.* **58** 13-30.
- Serfling, R.J. (1980). *Approximation Theorems of Mathematical Statistics*. Wiley, New York.
- Silverman, B.W. and Brown, T. (1978). Short distances, flat triangles and Poisson limits. *J. Appl. Prob.* **15** 815-825.
- Silverman, B.W. (1983). Convergence of a class of empirical distribution functions of dependent random variables. *Ann. Probab.* **11** 745-751.