# $L$-statistics for Repeated Measurements Data With Application to Trimmed Means, Quantiles and Tolerance Intervals

**Houssein I. Assaad**

Department of Statistics

Texas A&M University

College Station, TX 77843-3143, USA

**Pankaj K. Choudhary***

Department of Mathematical Sciences

University of Texas at Dallas

Richardson, TX 75083-0688, USA

**Abstract**

The $L$-statistics form an important class of estimators in nonparametric statistics. Its members include trimmed means and sample quantiles and functions thereof. This article is devoted to theory and applications of $L$-statistics for repeated measurements data, wherein the measurements on the same subject are dependent and the measurements from different subjects are independent. This article has three main goals: (a) Show that the $L$-statistics are asymptotically normal for repeated measurements data. (b) Present three statistical applications of this result, namely, location estimation using trimmed means, quantile estimation and construction of tolerance intervals. (c) Obtain a Bahadur representation for sample quantiles. These results are generalizations of similar results for independently and identically distributed data. The practical usefulness of these results is illustrated by analyzing a real data set involving measurement of systolic blood pressure. The properties of the proposed point and interval estimators are examined via simulation.

**Keywords and phrases**: Bahadur representation, Hadamard derivative, L-estimators, Nonparametric inference, Order statistics, Statistical functional, Weighted empirical process.

*Corresponding author. Email: pankaj@utdallas.edu, Tel: (972) 883-4436, Fax: (972)-883-6622.

## 1 Preliminaries

The class of linear functions of order statistics, the so-called $L$-statistics, plays a significant role in nonparametric statistics. Two prominent members of this class are sample quantiles and trimmed means. The sample quantiles are used for nonparametric estimation of population quantiles and their functions such as the inter-quartile range (David and Nagaraja 2003). They are also used for construction of nonparametric tolerance intervals for a population that are often sought in engineering, manufacturing and medicine (Krishnamoorthy and Mathew 2009). The trimmed means provide a robust alternative to sample mean for estimating a location parameter (Wilcox 2012). There is extensive literature on $L$-statistics for independently and identically distributed (i.i.d.) data — see, e.g., Serfling (1980), Huber (1981) and Fernholz (1983) for reviews. In particular, it is well-known that $L$-statistics are asymptotically normal for i.i.d. data. This article is concerned with generalizing this result to repeated measurements data and applying it to nonparametric estimation of trimmed means and quantiles, and construction of nonparametric tolerance intervals.

Let $X_{ij}$, $j = 1, \ldots, k_i$, denote the $k_i$ repeated measurements on the $i$th subject in the study, $i = 1, \ldots, n$. The subjects are assumed to be independent. The design of the study need not be balanced, i.e., the $k_i$ may not be equal. Let $N = \sum_{i=1}^{n} k_i$ denote the total number of observations. We also assume that:

A.1 The $X_{ij}$ are identically distributed as a continuous random variable $X$ with cumulative distribution function (c.d.f.) $F$, probability density $f$ and finite variance.

A.2 The vectors $(X_{i1}, \ldots, X_{ik_i})$, $i = 1, \ldots, n$, are independent, and $\exists$ an exchangeable sequence $\tilde{X}_1, \tilde{X}_2, \ldots$ such that $(X_{i1}, \ldots, X_{ik_i}) \overset{d}{=} (\tilde{X}_1, \ldots, \tilde{X}_{k_i})$ for each $i$. Thus, in particular, $\tilde{X}_1$ and $\tilde{X}_2$ represent two repeated measurements on a randomly selected subject from the population. Let $G$ be the bivariate c.d.f. of $(\tilde{X}_1, \tilde{X}_2)$. Due to the exchangeability assumption, both the marginals of $G$ are equal to $F$.

The distinctive feature of the data $X_{ij}$ is that the repeated measurements on a subject are replications of the same underlying quantity. In other words, the true underlying measurement for a subject does not change during the replication process. Therefore, the measurements on the same subject are dependent. On the other hand, the measurements from different subjects are independent. This kind of repeated measurements data are common in a variety of applications, including clinical studies concerned with estimation of reliability (Fleiss 1986; Dunn 1989), gauge repeatability and reliability studies (Burdick et al. 2005) and method comparison studies (Bland and Altman 1999). These data are typically analyzed by modeling them using a one-way random-effects model (or more generally a mixed-effects model) that treats the effect of subject as a random effect and assumes normal distributions for random effects and errors. The parameters of the model are estimated by a likelihood-based method and the asymptotic theory of maximum likelihood estimators (MLEs) is used for inference (Pinheiro and Bates 2000). Specifically, for inference on quantiles and construction of tolerance intervals, the methodology described in Krishnamoorthy and Mathew (2009, chap. 4) can be used (see also Sharma and Mathew (2012)). However, the MLEs are well-known to be non-robust against the violation of the normality assumption. This violation occurs frequently in practice — see Section 7 for a real example involving measurement of systolic blood pressure that motivated this work.

When the normality assumption is not reasonable, an alternative is to use a nonparametric method to analyze the data. Olsson and Rootzen (1996) consider nonparametric estimation of quantiles from repeated measurements. Their method can deal with unbalanced as well as balanced designs. Hutson (2003) considers nonparametric estimation of normal range — a quantile interval — using repeated measurements from a balanced design. These authors show that it is not a good idea to apply the estimation methods designed for i.i.d. data to univariate summaries of within-subject repeated measurements (e.g., averages) because it may lead to substantial loss of efficiency. The authors such as Wilcox (1994), Wilcox et al. (2000) and Keselman et al. (2000) use trimmed means in repeated measures designs in place of the usual means to get robust tests of hypotheses on treatment effects in an analysis-of-variance setting. Although the estimators studied in each of these articles are special cases of $L$-statistics, their authors are not concerned with studying the general class of $L$-statistics, as we do in this article. A study of general $L$-statistics allows us to present a unified treatment of the separate estimators. This unified approach additionally provides a method for constructing nonparametric tolerance intervals with repeated measurements data (see Section 5).

To study general $L$-statistics for repeated measurements data, let $X_{(1)} \leq X_{(2)} \leq \ldots \leq X_{(N)}$ be the order statistics associated with the $N$ observations $X_{ij}$, $j = 1, \ldots, k_i$, $i = 1, \ldots, n$. We estimate the population c.d.f. $F(x)$ by a weighted empirical c.d.f.,

$$F_n(x) = \sum_{i=1}^{n} w_i \sum_{j=1}^{k_i} I(X_{ij} \leq x), \tag{1}$$

where $w_i = w(k_i, n)$, $0 < w_i < 1$, is the known weight of the observation $X_{ij}$ and $I(A)$ is the indicator of event $A$. The weights depend on subject $i$ only through $k_i$ — the number of repeated measurements on the subject. All observations on a given subject receive the same weight because they are exchangeable from assumption A.2. The weights are assumed to satisfy $\sum_{i=1}^{n} k_i w_i = 1$, so that $F_n(x)$ is an unbiased estimator of $F(x)$.

The weights in $F_n$ may be arbitrary provided they satisfy the additional assumptions A.5 and A.6 in Section 2. In particular, these assumptions hold for the two weight functions,

$$w_{i,1} = \frac{1}{nk_i} \text{ and } w_{i,2} = \frac{1}{N}, \quad i = 1, \ldots, n, \tag{2}$$

which are of special interest due to their simplicity. The first one assigns a total of $1/n$ weight to each subject and distributes it equally among the repeated measurements on this subject, whereas the second one assigns equal weight to each observation in the data.

It can be seen that $F_n(x)$ is the minimum variance unbiased estimator of $F(x)$ if the weights are

$$\frac{\{1+(k_i-1)\rho(x,x)\}^{-1}}{\sum_{l=1}^{n} k_l\{1+(k_l-1)\rho(x,x)\}^{-1}}, \tag{3}$$

where

$$\rho(x,y) = \mathrm{corr}[I(\tilde{X}_1 \leq x), I(\tilde{X}_2 \leq y)] = \frac{G(x,y) - F(x)F(y)}{[F(x)\{1-F(x)\}F(y)\{1-F(y)\}]^{1/2}}. \tag{4}$$

Olsson and Rootzen (1996) refer to (3) as the "optimal weight function." The two weight functions in (2) are its special cases obtained by taking $\rho(x,x) = 1$ and $\rho(x,x) = 0$, respectively. All these weight functions are identical for balanced designs. We do not use the optimal weight function in this article as the resulting $F_n(x)$ is not a non-decreasing function of $x$ and the unknown $\rho(x,x)$ needs to be replaced with an estimate. These issues cause additional complications for the theory, but the optimal weights generally do not lead to significant gains in efficiency over the simpler weights in (2), especially $w_{i,1}$ (see, e.g., Olsson and Rootzen (1996)).

Next, for a given $0 < p < 1$, let $F_n^{-1}(p) = \inf\{x : F_n(x) \geq p\}$ denote the plug-in estimator of $F^{-1}(p) = \inf\{x : F(x) \geq p\}$, the $p$th quantile (or 100$p$th percentile) of the population. If we let $q_{s,N}$ be the total empirical probability weight of the $s$ smallest observations, then the order statistics are seen to be sample quantiles, i.e.,

$$X_{(s)} = F_n^{-1}(p), \text{ if } q_{s-1,N} < p \leq q_{s,N}, \; s = 1, \ldots, N. \tag{5}$$

Here $q_{0,N} = 0$ and $q_{N,N} = 1$.

A general $L$-statistic has the form:

$$\sum_{s=1}^{N} c_{s,N} X_{(s)}, \tag{6}$$

for some choice of constants $c_{1,N}, \ldots, c_{N,N}$. Consider a fixed signed measure $dM(x) = m(x)dx$ on $[0,1]$. The function $m(x)$ is sometimes called a *weight-generating function*. An important subclass of (6) wide enough for all typical applications is given by Serfling (1980, chap. 8):

$$T(F_n) = \int_0^1 F_n^{-1}(x)m(x)dx + \sum_{l=1}^{r} a_l F_n^{-1}(p_l) =: T_1(F_n) + T_2(F_n), \tag{7}$$

for a pre-specified positive integer $r$. It is also assumed that $0 < p_1 < p_2 < \ldots < p_r < 1$ are specified, and that $a_1, \ldots, a_r$ are known constants, not all of which are equal to zero. The statistic $T(F_n)$ can be written in the more familiar $L$-statistic form (6) by using (5) and taking the coefficients as $c_{s,N} = \int_{q_{s-1,N}}^{q_{s,N}} m(x)dx + a_l$, where $l$ is such that $q_{s-1,N} < p_l \leq q_{s,N}$. The form (7) shows that $T(F_n)$ is actually a sum of two $L$-statistics: $T_1(F_n)$ — the *continuous part* of $T(F_n)$, obtained by weighting all observations in a continuous manner; and $T_2(F_n)$ — the *discrete part* of $T(F_n)$, which is a weighted sum of $r$ observations. Often, the statistic of interest is $T_1(F_n)$ alone (e.g., $100\alpha\%$ trimmed mean, $0 < \alpha < 1/2$) or $T_2(F_n)$ alone (e.g., sample quantile). Upon replacing $F_n$ in (7) with $F$, we get the $L$-functional,

$$T(F) = \int_0^1 F^{-1}(x)m(x)dx + \sum_{l=1}^{r} a_l F^{-1}(p_l) =: T_1(F) + T_2(F), \tag{8}$$

representing the population parameter that $T(F_n)$ actually estimates.

In the standard i.i.d. case, which is a special case of our set up when $k_i = 1 \;\forall\, i$, the statistic $T$ is known to be asymptotically normal (Serfling 1980, p. 282). We generalize this result to the case of repeated measurements data in Section 2 using a *statistical functional* approach (van der Vaart 1998, chap. 20). In particular, we extend the technique of Fernholz (1983, prop. 4.3.3) for i.i.d. data to show that the remainder term in the von Mises expansion of a Hadamard differentiable functional goes to zero in probability. Further, we extend the technique of Ghosh (1971, thm. 1) for the i.i.d. data to get a Bahadur representation for sample quantiles. Together these results provide the desired asymptotic normality of $T$. This result is applied in Sections 3, 4 and 5 respectively for location estimation using trimmed means, quantile estimation and construction of tolerance intervals. We perform a simulation study in Section 6 to examine the finite sample accuracy of the proposed confidence and tolerance intervals and also to compare the two weight functions in (2). A real data application is presented in Section 7. Section 8 is devoted to technical details.

# 2 Asymptotic normality of $T(F_n)$

First, we make the following assumptions in addition to A.1 and A.2.

A.3 $\max_{i=1,\ldots,n} k_i \leq k^*$, where $k^*$ is a known constant. Thus, in the asymptotic analysis, we let the number of subjects increase but keep the number of repeated measurements bounded.

A.4 Let $\mu_n(k)$ denote the proportion of subjects with exactly $k$ repeated measurements, $k = 1, \ldots, k^*$. There exist constants $\mu(k)$ such that $\lim_{n\to\infty} \mu_n(k) = \mu(k)$, $k = 1, \ldots, k^*$. If for some $k$, there is no subject in the study with $k$ measurements, then $\mu_n(k)$ and $\mu(k)$ are simply ignored. So, without loss of generality, $\mu_n(k)$ and $\mu(k)$, $k = 1, \ldots, k^*$, are all taken to be positive.

A.5 Let $w(k) = w(k, n)$, $k = 1, \ldots, k^*$, denote the common weight of observations from subjects with $k$

repeated measurements. There exist constants $\theta(k)$ such that $\lim_{n\to\infty} nw(k) = \theta(k)$, $k = 1, \ldots, k^*$. As in A.4, the $\theta(k)$ are assumed to be positive without loss of generality.

A.6 The ratio $(\max_{1\le i\le n} w_i)\,/\,(\min_{1\le i\le n} w_i) = o\big(n^{\delta/\{2(2+\delta)\}}\big)$ for some $\delta > 0$.

A.7 The function $m(x)$ has support in $[\gamma, 1-\gamma]$ for some $0 < \gamma < 1/2$, and $\exists\, C > 0$ such that $|m(x)| \le C$.

Let $IC(x, F, T)$ denote the *influence curve* of the functional $T$ in (8). It is defined as: $IC(x, F, T) = (d/dt)T(F + t(\delta_x - F))|_{t=0}$, where $\delta_x$ is the c.d.f. of the point mass distribution at $x$ (van der Vaart 1998, chap. 20). In other words, $\delta_x(y) = I(x \le y)$, $y \in \mathbb{R}$. Since $T(F) = T_1(F) + T_2(F)$, we can write

$$IC(x, F, T) = IC(x, F, T_1) + IC(x, F, T_2). \tag{9}$$

The influence curves for $T_1$ and $T_2$ have been derived, for instance, in Huber (1981, pp. 56-57). They are:

$$IC(x, F, T_1) = \int_{-\infty}^{x} m(F(y))dy - \int_{-\infty}^{+\infty} (1 - F(y))m(F(y))dy, \tag{10}$$

$$IC(x, F, T_2) = \sum_{l=1}^{r} a_l \frac{p_l - \delta_x(F^{-1}(p_l))}{f(F^{-1}(p_l))}. \tag{11}$$

Next, let $\psi^2(k_i) = \mathrm{var}[\sum_{j=1}^{k_i} IC(X_{ij}, F, T)]$. Due to the exchangeability assumption A.2, we can write

$$\begin{aligned}\psi^2(k_i) &= k_i \mathrm{var}[IC(\tilde{X}_1, F, T)] + 2\binom{k_i}{2}\mathrm{cov}[IC(\tilde{X}_1, F, T), IC(\tilde{X}_2, F, T)] \\ &= k_i E[IC^2(\tilde{X}_1, F, T)] + 2\binom{k_i}{2} E[IC(\tilde{X}_1, F, T) IC(\tilde{X}_2, F, T)],\end{aligned} \tag{12}$$

where the second equality follows from the fact that an influence curve has mean zero (van der Vaart 1998, chap. 20). Also, let

$$\sigma_n^2 = n\sum_{i=1}^{n} w_i^2 \psi^2(k_i) = \sum_{k=1}^{k^*} \mu_n(k)\{nw(k)\}^2\psi^2(k), \quad \sigma^2 = \lim_{n\to\infty} \sigma_n^2 = \sum_{k=1}^{k^*} \mu(k)\theta^2(k)\psi^2(k), \tag{13}$$

where $k^*$, $w(k)$, $\mu_n$, $\mu$ and $\theta$ are from A.4 and A.5. The following result gives asymptotic normality of $T(F_n)$.

**Theorem 1.** *Consider the L-functional $T$ defined in (8) and $\sigma^2$ defined in (13). Then, under the assumptions A.1 to A.7, and the additional assumptions listed in Proposition 3 in Section 8, we have:*

$$n^{1/2}[T(F_n) - T(F)] \xrightarrow{d} N(0, \sigma^2).$$

This result generalizes a similar result in Serfling (1980, thm. A, p. 282) for i.i.d. data. It can be used in the usual manner to perform large-sample inference on $T(F)$. For example, when $n$ is large, an approximate $100(1-\beta)\%$ confidence interval for $T(F)$ is:

$$T(F_n) \pm z_{1-\beta/2}\,\hat{\sigma}_n/n^{1/2}, \tag{14}$$

where $\hat{\sigma}_n^2$ is an estimator of $\sigma_n^2$ whose limit is $\sigma^2$, and $z_{1-\beta/2}$ is the $(1-\beta/2)$th quantile of the $N(0,1)$ distribution. A general approach to get $\hat{\sigma}_n^2$ is to simply replace the population quantities in $\sigma_n^2$ with their sample counterparts. In particular, let $\widehat{IC}_{ij} = IC(X_{ij}, F_n, T)$ denote the *empirical influence curve*, obtained by replacing $F$ in (9) with $F_n$. Then the expectations needed in (12) can be estimated as:

$$\begin{aligned}
\widehat{E}[IC^2(\tilde{X}_1, F, T)] &= \frac{1}{n}\sum_{i=1}^{n}\frac{1}{k_i}\sum_{j=1}^{k_i}\widehat{IC}_{ij}^2, \\
\widehat{E}[IC(\tilde{X}_1, F, T)IC(\tilde{X}_2, F, T)] &= \frac{1}{\#\{i : k_i > 1\}}\sum_{i:k_i>1}\frac{1}{k_i(k_i-1)}\sum_{1\le j\ne l\le k_i}\widehat{IC}_{ij}\widehat{IC}_{il}.
\end{aligned}$$

Plugging-in these estimates in (12) gives $\hat{\psi}^2(k_i)$. Hence from (13), $\hat{\sigma}_n^2$ can be taken as $\hat{\sigma}_n^2 = n\sum_{i=1}^{n} w_i^2\hat{\psi}^2(k_i)$. Often, however, the expression for $\sigma^2$ can be simplified (see, e.g., (17) in Section 4). In this case, the unknowns in the simplified expression may be replaced with their estimates to get $\hat{\sigma}_n^2$.

## 3 Estimation of population trimmed means

For a given $0 < \alpha < 1/2$, the $100\alpha\%$ trimmed mean can be obtained from the functional $T$ in (8) by taking $m(x) = I(\alpha < x < 1-\alpha)/(1-2\alpha)$ in its continuous part $T_1$, and setting its discrete part $T_2$ equal to zero. This gives the population and sample versions of the trimmed mean as:

$$T(F) = \frac{1}{1-2\alpha}\int_{\alpha}^{1-\alpha} F^{-1}(x)dx, \quad T(F_n) = \frac{1}{1-2\alpha}\int_{\alpha}^{1-\alpha} F_n^{-1}(x)dx.$$

Here $\alpha$ is the trimming proportion on each side. To write the sample version in the familiar $L$-statistic form, let $l^*$ and $u^*$ be integers such that $q_{l^*-1,N} < \alpha \le q_{l^*,N}$ and $q_{u^*-1,N} < 1-\alpha \le q_{u^*,N}$. Also, let $w_j^*$, an element of $\{w_i,\, i = 1, ..., N\}$ in (1), be the weight associated with the $j$th order statistic $X_{(j)}$. Then, from (5), the sample $\alpha$-trimmed mean is:

$$T(F_n) = \frac{1}{1-2\alpha}\big\{(q_{l^*,N}-\alpha)X_{(l^*)} + \sum_{j=l^*+1}^{u^*-1} w_j^* X_{(j)} + (1-\alpha-q_{u^*-1,N})X_{(u^*)}\big\}. \tag{15}$$

It may be noted that $T(F)$ coincides with the median if the distribution $F$ is symmetric. Huber (1981) gives the influence function of $T(F)$ as:

$$IC(X,F,T) = \begin{cases} \frac{1}{1-2\alpha}\{F^{-1}(\alpha) - W(F)\}, & \text{if } X < F^{-1}(\alpha), \\ \frac{1}{1-2\alpha}\{X - W(F)\}, & \text{if } F^{-1}(\alpha) \le X \le F^{-1}(1-\alpha), \\ \frac{1}{1-2\alpha}\{F^{-1}(1-\alpha) - W(F)\}, & \text{if } X > F^{-1}(1-\alpha), \end{cases}$$

where $W(F) = (1-2\alpha)T(F) + \alpha\{F^{-1}(\alpha) + F^{-1}(1-\alpha)\}$. This influence curve can be used as described in Section 2 to estimate the standard error of the sample trimmed mean and to get an approximate confidence interval for the population trimmed mean.

# 4 Estimation of population quantiles

For a given $0 < p < 1$, the $p$th quantile is a special case of the functional $T$ in (8) obtained by setting its continuous part $T_1$ equal to zero, and taking $r = 1$, $p_1 = p$ and $a_1 = 1$ in its discrete part $T_2$. This gives the population and the sample $p$th quantile as $T(F) = F^{-1}(p)$ and $T(F_n) = F_n^{-1}(p)$. These quantities will henceforth be denoted as $Q_p$ and $\hat{Q}_p$ for notational convenience. From (11), the influence curve for $Q_p$ is:

$$IC(X,F,Q_p) = \{p - \delta_X(Q_p)\}/f(Q_p). \tag{16}$$

Upon substituting this expression in (12) and simplifying (13), we get:

$$\sigma^2 = \frac{p(1-p)}{f^2(Q_p)} \sum_{k=1}^{k^*} k\{1 + (k-1)\rho(Q_p, Q_p)\}\, \mu(k)\, \theta^2(k), \tag{17}$$

where $\rho(Q_p, Q_p)$ is given by (4). The asymptotic normality of $\hat{Q}_p$ holds from Theorem 1.

It may be noted that Olsson and Rootzen (1996) also establish asymptotic normality of $\hat{Q}_p$ by using in $F_n$ an estimate of the optimal weight function (3), obtained by replacing $\rho(x,x)$ with an estimator $\hat{\rho}(x,x)$ which is to be defined in (19). Although the weights in our result do not depend on $x$, they can be arbitrary provided they satisfy the assumptions A.5 and A.6. In this sense, our result differs from Olsson and Rootzen's. Besides, unlike theirs, our result follows from a more general result derived for $L$-statistics.

Using (13), $\sigma^2$ given in (17) can be estimated by

$$\hat{\sigma}_n^2 = \frac{np(1-p)}{\hat{f}^2(\hat{Q}_p)} \sum_{i=1}^{n} k_i\{1 + (k_i - 1)\hat{\rho}(\hat{Q}_p, \hat{Q}_p)\}\, w_i^2, \tag{18}$$

where $\hat{f}$ is an estimator of the density $f$ and $\hat{\rho}$ is an estimator of $\rho$. The density may be estimated as

$$\hat{f}(x) = \frac{F_n(x+h) - F_n(x-h)}{2h},$$

with the bandwidth $h$ chosen, e.g., according to the recommendations of Silverman (1986, chap. 4). Next, the correlation $\rho(x, y)$ may be estimated by a simple estimator,

$$\hat{\rho}(x,y) = \frac{\widehat{cov}\{I(\tilde{X}_1 \le x), I(\tilde{X}_2 \le y)\}}{\left\{\widehat{var}[I(\tilde{X}_1 \le x)]\widehat{var}[I(\tilde{X}_1 \le y)]\right\}^{1/2}}, \tag{19}$$

where

$$\widehat{var}[I(\tilde{X}_1 \le x)] = \frac{1}{\#\{i : k_i > 1\}} \sum_{i:k_i>1} \frac{1}{k_i} \sum_{j=1}^{k_i} \left\{I(X_{ij} \le x) - \overline{F}_n(x)\right\}^2,$$

$$\widehat{cov}[I(\tilde{X}_1 \le x), I(\tilde{X}_2 \le y)] = \frac{1}{\#\{i : k_i > 1\}} \sum_{i:k_i>1} \frac{1}{k_i(k_i-1)} \times \sum_{1\le j \ne l \le k_i} \{I(X_{ij} \le x) - \overline{F}_n(x)\}\{I(X_{il} \le y) - \overline{F}_n(y)\},$$

$$\overline{F}_n(x) = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{k_i} \sum_{j=1}^{k_i} I(X_{ij} \le x).$$

This $\hat{\rho}$ is related to the estimator of an intraclass correlation given in Karlin et al. (1981) and has also been used in Olsson and Rootzen (1996). Potentially other estimators of $\rho$ may also be considered, but $\hat{\rho}$ works well in simulations.

Although $\hat{\sigma}_n^2$ defined in (18) can be used in (14) to get a confidence interval for $Q_p$, it has the unattractive feature of having to estimate the density $f(Q_p)$. This problem can be avoided by constructing the confidence interval directly using the following result. It is proved in Section 8.4.

**Theorem 2.** *Suppose the assumptions A.1 to A.6 hold. Assume also that the bivariate c.d.f. $G$ of $(\tilde{X}_1, \tilde{X}_2)$ is continuous at $(Q_p, Q_p)$ and $f(Q_p) > 0$, for $0 < p < 1$. Let $\hat{r}_n^2 = np(1-p)\sum_{i=1}^{n} k_i\{1 + (k_i - 1)\hat{\rho}(\hat{Q}_p, \hat{Q}_p)\}\, w_i^2$. Define $\hat{l}_n = p - z_{1-\beta/2}\hat{r}_n/n^{1/2}$ and $\hat{u}_n = p + z_{1-\beta/2}\hat{r}_n/n^{1/2}$. Then, $\lim_{n\to\infty} P(Q_p \in [\hat{Q}_{\hat{l}_n}, \hat{Q}_{\hat{u}_n}]) = 1 - \beta$.*

We next obtain a weak version of Bahadur representation of sample quantiles which generalizes Ghosh (1971, thm. 1) for i.i.d. data. It is proved in Section 8.5.

**Theorem 3.** *Suppose the assumptions A.1 to A.6 hold. Assume also that the bivariate c.d.f. $G$ of $(\tilde{X}_1, \tilde{X}_2)$ is continuous at $(Q_p, Q_p)$ and $f(Q_p) > 0$, for $0 < p < 1$. Let $p^{(n)}$ be a sequence of probabilities such that*

$n^{1/2}\left(p^{(n)}-p\right)=O(1)$, *and* $\hat{Q}_{p^{(n)}}=F_n^{-1}\left(p^{(n)}\right)$. *Then,*

$$\hat{Q}_{p^{(n)}}=Q_p+\{p^{(n)}-F_n(Q_p)\}/f(Q_p)+o_p(1/n^{1/2}).$$

# 5 Construction of nonparametric tolerance intervals

For given $0<p,\beta<1$, an interval $[\hat{L}_1,\hat{L}_2]$ computed from the sample data is called a $(p,1-\beta)$ *tolerance interval* for a random variable $X$ if

$$P(F(\hat{L}_2)-F(\hat{L}_1)\geq p)=1-\beta. \tag{20}$$

The random quantity $F(\hat{L}_2)-F(\hat{L}_1)=\int_{\hat{L}_1}^{\hat{L}_2}f(x)dx$ represents the *probability content* of the interval $[\hat{L}_1,\hat{L}_2]$ under the distribution of $X$. Thus, a $(p,1-\beta)$ tolerance interval captures at least $p$ proportion of the $X$ population with $1-\beta$ confidence. The interval is one-sided if either $\hat{L}_1=-\infty$ or $\hat{L}_2=\infty$, otherwise it is two-sided. Tolerance intervals are common in engineering and manufacturing applications; see Guttman (1988), Vardeman (1992) and Krishnamoorthy and Mathew (2009) for an introduction to this topic.

In general, a nonparametric tolerance interval has the form $[\hat{L}_1,\hat{L}_2]=[X_{(r)},X_{(s)}]$, where $r$ and $s$ $(r<s)$ are chosen to satisfy (20). This notation allows the possibility of one-sided intervals by letting $r$ be zero with $X_{(0)}=-\infty$ and $s$ be $N+1$ with $X_{(N+1)}=\infty$, provided both $r=0$ and $s=N+1$ are not taken at the same time. In the i.i.d. case, it is well-known that $F\{X_{(s)}\}-F\{X_{(r)}\}$ follows a Beta $(s-r,\,N-s+r+1)$ distribution (Guttman, 1988). Hence, for example, a two-sided equal-tailed tolerance interval can be obtained by taking $s=N-r+1$, $r<(N+1)/2$, and numerically solving (20) for $r$. This is equivalent to finding $r$ such that the c.d.f. of the Beta $(N-2r+1,2r)$ distribution at $p$ is $\beta$.

In the case of repeated measurements data, however, the distribution of $F\{X_{(s)}\}-F\{X_{(r)}\}$ (or equivalently $F(\hat{Q}_{p_2})-F(\hat{Q}_{p_1})$ for $p_2>p_1$) does not have a simple form. This motivates us to search for $p_1$ and $p_2$ so that (20) holds in the limit, i.e.,

$$\lim_{n\to\infty}P(F(\hat{Q}_{p_2})-F(\hat{Q}_{p_1})\geq p)=1-\beta. \tag{21}$$

We refer to the resulting $(\hat{Q}_{p_1},\hat{Q}_{p_2})$ as an *asymptotic* $(p,1-\beta)$ tolerance interval. This interval has approximately $1-\beta$ confidence when $n$ is large. As before, here we allow the possibility of one-sided intervals by letting $p_1$ be zero with $\hat{Q}_0=-\infty$ and $p_2$ be one with $\hat{Q}_1=\infty$, provided both $p_1=0$ and $p_2=1$ are not taken simultaneously. To develop a procedure for constructing this interval, let

$$\nu_l^2(p_l) = p_l(1-p_l)\sum_{k=1}^{k^*} k\left\{1+(k-1)\rho(Q_{p_l},Q_{p_l})\right\}\mu(k)\theta^2(k), \quad l=1,2,$$
$$\nu_{12}(p_1,p_2) = p_1(1-p_2)\sum_{k=1}^{k^*} k\left[1+(k-1)\rho(Q_{p_1},Q_{p_2})\left\{\frac{(1-p_1)p_2}{p_1(1-p_2)}\right\}^{1/2}\right]\mu(k)\theta^2(k),$$
$$\nu^2(p_1,p_2) = \nu_1^2(p_1) - 2\nu_{12}(p_1,p_2) + \nu_2^2(p_2), \tag{22}$$

where $\rho(x,y)$ is given by (4). Here $\nu_l^2(p_l)$ is defined for $0<p_l<1$, and $\nu_{12}(p_1,p_2)$ and $\nu^2(p_1,p_2)$ are defined for $0<p_1<p_2<1$. Next, let $\hat{\nu}_1^2$, $\hat{\nu}_2^2$ and $\hat{\nu}^2$ be consistent estimators of $\nu_1^2$, $\nu_2^2$ and $\nu^2$, respectively. They are obtained by replacing $Q_p$ and $\rho$ in (22) with $\hat{Q}_p$ and $\hat{\rho}$, defined by (19). The next result shows that the probability content of $(\hat{Q}_{p_1},\hat{Q}_{p_2})$ is asymptotically normal regardless of whether this interval is one- or two-sided. It is a consequence of Theorem 1 and is proved in Section 8.6.

**Theorem 4.** *Suppose that the assumptions A.1 to A.6 hold.*

*(a) Suppose that the assumptions listed in Proposition 3 in Section 8 also hold for $r=2$ and for all $0<p_1<p_2<1$. Then, $n^{1/2}[F(\hat{Q}_{p_2})-F(\hat{Q}_{p_1})-(p_2-p_1)] \xrightarrow{d} N(0,\nu^2(p_1,p_2))$.*

*(b) Suppose that the assumptions listed in Proposition 3 also hold for $r=1$ and for all $0<p_l<1$, separately for each $l=1,2$. Then, $n^{1/2}(F(\hat{Q}_{p_l})-p_l) \xrightarrow{d} N(0,\nu_l^2(p_l))$, $l=1,2$.*

This result implies that, when $n$ is large, $F(\hat{Q}_{p_2})-F(\hat{Q}_{p_1})$ and $F(\hat{Q}_{p_l})$ approximately follow $N(p_2-p_1,\hat{\nu}^2(p_1,p_2)/n)$ and $N(p_l,\hat{\nu}_l^2(p_l)/n)$ distributions, respectively. Therefore, $p_1$ and $p_2$ required for the two-sided interval $(\hat{Q}_{p_1},\hat{Q}_{p_2})$ to satisfy (21) can be found by solving:

$$n^{1/2}\{p-(p_2-p_1)\}/\hat{\nu}(p_1,p_2) = z_\beta. \tag{23}$$

It follows from (23) that $p_1$ and $p_2$ satisfy $p_2-p_1\geq p$ whenever $0<\beta\leq 1/2$. For an equal-tailed interval one can take $p_2=1-p_1$ in (23). Analogously, for the one-sided case, $p_1$ needed for the interval $(\hat{Q}_{p_1},\infty)$ and $p_2$ needed for the interval $(-\infty,\hat{Q}_{p_2})$ can be computed by respectively solving the equations

$$n^{1/2}\{p-(1-p_1)\}/\hat{\nu}_1(p_1) = z_\beta, \quad n^{1/2}(p-p_2)/\hat{\nu}_2(p_2) = z_\beta. \tag{24}$$

The finite sample accuracy of these tolerance intervals can be improved by computing $(p_1,p_2)$ after applying a logit (or log-odds) transformation to the probability content. For this, we can deduce from

Theorem 4 and delta method that:

$$n^{1/2}[\text{logit}\{F(\hat{Q}_{p_2}) - F(\hat{Q}_{p_1})\} - \text{logit}(p_2 - p_1)] \xrightarrow{d} N(0,\, \nu^2(p_1, p_2)/\{(p_2 - p_1)(1 - p_2 + p_1)\}^2),$$
$$n^{1/2}[\text{logit}\{F(\hat{Q}_{p_l})\} - \text{logit}(p_l)] \xrightarrow{d} N(0, \nu_l^2(p_l)/\{p_l(1 - p_l)\}^2), \quad l = 1, 2.$$

Thus, the more accurate $(p_1, p_2)$ can be computed by solving the following counterparts of (23) and (24):

$$n^{1/2}\{\text{logit}(p) - \text{logit}(p_2 - p_1)\}(p_2 - p_1)(1 - p_2 + p_1)/\hat{\nu}(p_1, p_2) = z_\beta,$$
$$n^{1/2}\{\text{logit}(p) - \text{logit}(1 - p_1)\}p_1(1 - p_1)/\hat{\nu}_1(p_1) = z_\beta, \quad n^{1/2}\{\text{logit}(p) - \text{logit}(p_2)\}p_2(1 - p_2)/\hat{\nu}_2(p_2) = z_\beta. \tag{25}$$

This is the method we recommend for use in practice.

# 6 A simulation study

In this section, we use Monte Carlo simulations to evaluate certain properties of sample trimmed means, sample quantiles and tolerance intervals. We also compare the two weight functions given in (2) for estimating $F$ in the case of unbalanced designs (recall that they are equal in the case of balanced designs). Our focus is on models that have the structure of a one-way random-effects model:

$$X_{ij} = \xi + 3b_i + \epsilon_{ij}, \quad j = 1, \ldots, k_i, \quad i = 1, \ldots, n, \tag{26}$$

where $\xi$ is the fixed intercept taken to equal 0 without loss of generality, $b_i$ is the random effect of the $i$th subject and $\epsilon_{ij}$ is the random error term. Here $b_i$ and $\epsilon_{ij}$ are mutually independent and they are also independent for different subjects. The coefficient of $b_i$ in (26) is taken as 3 to have high intraclass correlation between the repeated measurements, which is a typical scenario in applications.

## 6.1 Trimmed Means

We first examine the asymptotic efficiency of the trimmed mean relative to the normality-based MLE of the underlying location parameter. We specifically consider a total of ten models obtained using combinations of $t_3$, $t_5$, $t_{30}$ and $N(0, 1)$ as distributions for the two random terms in (26). These models are summarized in Table 1. Only symmetric distributions are considered so that the parameter $T(F)$ that the trimmed mean estimates equals the location parameter $\xi$, whose true value is zero.

We simulate data from each model on $n = 400$ subjects in a way that $k_i$ equals 1, 2, 3 and 4 for

100 subjects each. These data are used to compute three estimates of $\xi$ — the $\alpha$-trimmed mean with weights $w_{i,1} = 1/(nk_i)$ and $w_{i,2} = 1/N$, and the MLE of $\xi$ assuming normality for both random-effects and errors in the model (26). These estimators are denoted as $\hat{\xi}_1$, $\hat{\xi}_2$ and $\hat{\xi}_{mle}$, respectively. Three values for $\alpha$ are used: 0.05, 0.10 and 0.125. The process of simulating data and estimating $\xi$ is repeated 2,000 times, and the approximate mean-squared errors (MSEs) of the three estimators are computed. The ratio MSE($\hat{\xi}_{mle}$)/MSE($\hat{\xi}_l$) gives the estimated asymptotic relative efficiency (ARE) of $\hat{\xi}_l$ relative to $\hat{\xi}_{mle}$, $l = 1, 2$. The computations are performed using the statistical software `R` (R Development Core Team 2011) and its `nlme` package (Pinheiro et al. 2011) is used to get $\hat{\xi}_{mle}$.

Table 1 presents the ARE estimates. It shows that $\hat{\xi}_1$ is more efficient than $\hat{\xi}_2$ at all settings considered. In fact, $\hat{\xi}_1$ is only slightly less efficient than the MLE. In the worst case, $\hat{\xi}_1$ loses 5% efficiency over the MLE, which occurs when $\alpha = 0.125$ and the model is either $3N(0,1)+N(0,1)$ or $3N(0,1)+t_5$. On the other hand, the gain in efficiency of $\hat{\xi}_1$ over the MLE can be substantial in case of heavy-tailed distributions. The largest gain in the table is 85% for the model $3t_3 + t_5$ and $\alpha = 0.125$. It is also interesting to note that the heavy-tailedness of random-effect distribution causes more loss in efficiency of $\hat{\xi}_{mle}$ than the heavy-tailedness of error distribution. Moreover, when the model for error distribution is fixed and the random-effect distribution spans $N(0,1)$, $t_5$ and $t_3$ distributions, we observe the pattern that the ARE of $\hat{\xi}_1$ for $t_5$ falls between those for $N(0,1)$ and $t_3$. But this pattern does not hold when the random-effect distribution is fixed and the error distribution varies. Additional simulations in Assaad (2012, chap. 4) for balanced designs with between 2 to 4 repeated measurements per subject show that the above conclusion regarding the relative merits of $\hat{\xi}_1$ and $\hat{\xi}_{mle}$ remains unchanged. (It may be recalled that $\hat{\xi}_1 = \hat{\xi}_2$ in case of balanced designs.) Overall, these findings suggest that $\hat{\xi}_1$ with $\alpha = 0.10$ or 0.125 provides a strong alternative to $\hat{\xi}_{mle}$ in all models considered.

Next, we examine the coverage accuracy of two nonparametric confidence intervals for $\xi$ — one using $\hat{\xi}_1$ and the other using $\hat{\xi}_2$. Simulations in Assaad (2012, chap. 4) show that $n$ around 50 is large enough for these confidence intervals to be accurate. Moreover, just like the ARE case, the design of the study (balanced or unbalanced) and the number of repeated measurements per subject do not have any noteworthy impact on this conclusion.

## 6.2 Quantiles

Here we only evaluate the finite sample accuracy of the confidence interval for $Q_p$ obtained using Theorem 2. For a comparison of asymptotic efficiencies of $\hat{Q}_p$ with weights $w_{i,1} = 1/(nk_i)$ and $w_{i,2} = 1/N$, we refer the reader to Figure 1 of Olsson and Rootzen (1996). It shows that unless the correlation $\rho(Q_p, Q_p)$, given by (4), is small, $w_{i,1}$ leads to a more efficient estimator than $w_{i,2}$.

To study the coverage accuracy, we consider three distributions — $N(0,1)$, $t_3$ and a skew-normal distribution (Azzalini 1985) with location zero, scale one and skewness parameter 5, denoted as $SN(0,1,5)$ — for each of the two random terms $b_i$ and $\epsilon_{ij}$ in (26). This results in a total of nine models. From each model, we simulate data on $n = 52$ subjects in a way that $k_i$ equals $1,2,3,4$ for 13 subjects each in case of an unbalanced design and $k_i$ equals $2,3,4$ for all subjects in case of balanced designs. These data are used to compute 95% confidence intervals for median $Q_{0.5}$ and 90th percentile $Q_{0.9}$ via Theorem 2. Simulations in Assaad (2012, chap. 4) reveal that $n$ around 50 may be large enough for these confidence intervals to be accurate. Besides, this accuracy does not seem to be affected by either the design of the study (balanced or unbalanced) or the data distribution (normal, heavy-tailed or skewed) or the number of repeated measurements. Further simulations for $Q_{0.99}$ (not presented here) show that $n$ around 250 is needed to achieve satisfactory coverage probabilities in all the above models.

## 6.3 Tolerance intervals

Here we examine the finite sample accuracy of the proposed tolerance intervals. As in Section 6.2, we focus on nine models of the form (26). They are summarized in the first column of Table 2. From each model, we simulate data on $n = 60$ subjects in a way that $k_i$ equals $1,2,3,4$ for 15 subjects each. These data are used to compute two-sided equal-tailed tolerance intervals by solving (25), using each of the two weight functions $w_{i,1} = 1/(nk_i)$ and $w_{i,2} = 1/N$. We then compute the true probability content of each interval numerically. This process of simulating data, constructing tolerance intervals and computing their probability content is repeated 2,000 times and the proportion of times the true content exceeds $p$ is obtained.

Table 2 presents these proportions for $p = 0.80, 0.90$ and $1 - \beta = 0.95$. We see that, in general, the values are closer to 0.95 in case of $p = 0.80$ than $p = 0.90$, and with weights $w_{i,1}$ than $w_{i,2}$. Specifically with weights $w_{i,1}$, most values are around 0.94 in case of $p = 0.80$ and around 0.93 in case of $p = 0.90$, regardless of whether the distribution is normal, heavy-tailed or skewed. On the whole, these values that the tolerance intervals with weights $w_{i,1}$ have reasonable accuracy with $n = 60$ in case of $p = 0.80$, whereas a larger $n$ (around 80, based on additional simulations in Assaad (2012, chap. 4)) is needed to achieve a similar level of accuracy in case of $p = 0.90$. Further simulations for balanced designs with between 2 to 4 repeated measurements per subject suggest that the accuracy of the tolerance interval does not depend on the number of repeated measurements.

## 7 Application to blood pressure data

In this section, we use a portion of the blood pressure data of Bland and Altman (1999) to illustrate the application of our results. These data were originally collected to evaluate agreement between three methods of measuring systolic blood pressure. However, since a comparison of two or more measurement methods is not of concern in this article, we focus only on the data from one of the methods, namely, the semi-automatic blood pressure monitor. There are 3 repeated measurements (in mmHg) of systolic blood pressure taken using the monitor in quick succession on each of 85 subjects in the study. These measurements are our $X_{ij},\ j = 1, 2, 3,\ i = 1, \ldots, 85$, and $X$ represents the population from which these data are drawn. We are interested in estimating the center, the 90th and 99th percentiles and the 10% trimmed mean of the distribution of $X$, and also constructing a $(p = 0.90, 1 - \beta = 0.95)$ tolerance interval for it. A histogram of the data presented in Figure 1 shows marked right-skewness in the distribution.

We first fit a one-way random-effects model,

$$X_{ij} = \xi + b_i + \epsilon_{ij}, \quad j = 1, 2, 3, \quad i = 1, \ldots, 85, \tag{27}$$

assuming that $b_i \sim N(0, \sigma_b^2)$ and $\epsilon_{ij} \sim N(0, \sigma_\epsilon^2)$. This model implies that $X \sim N(\xi, \sigma_b^2 + \sigma_\epsilon^2)$. The model is fit using the `nlme` package (Pinheiro et al. 2011) in `R`. The MLE of the parameter vector $(\xi, \sigma_b^2, \sigma_\epsilon^2)$ and its approximate estimated variance matrix are:

$$\begin{pmatrix} 143.03 \\ 971.30 \\ 83.14 \end{pmatrix}, \quad \begin{pmatrix} 11.75 & 0.00 & 0.00 \\ 0.00 & 23492.14 & -27.11 \\ 0.00 & -27.11 & 81.32 \end{pmatrix}.$$

Figure 2 presents the normal quantile-quantile plot of the estimated random-effects and the residuals. There is clear evidence of skewness in random-effects and heavier-than-normal tails in residuals, invalidating the normality assumption and justifying the need for a nonparametric analysis.

Table 3 summarizes the ML and nonparametric estimates of median $Q_{0.5}$, 90th percentile $Q_{0.9}$ and 99th percentile $Q_{0.99}$, along with their standard errors and 95% confidence intervals. It may be noted that the two weight functions in (2) used for nonparametric estimation are identical due to the balanced design of the data. In the parametric case, $Q_p = \xi + z_p(\sigma_b^2 + \sigma_\epsilon^2)^{1/2}$ and $\hat{Q}_p = \hat{\xi} + z_p(\hat{\sigma}_b^2 + \hat{\sigma}_\epsilon^2)^{1/2}$ is its MLE. Further, the delta method (Lehmann 1999, p. 295) is used to estimate the standard error of $\hat{Q}_p$ and to construct the confidence interval for $Q_p$. In the nonparametric case, the standard error of $\hat{Q}_p$ is estimated using (18), with $h = 0.79(\hat{Q}_{0.75} - \hat{Q}_{0.25})n^{-1/5}$ as the bandwidth in the density estimate $\hat{f}$ (Silverman 1986,

p. 47), and the confidence interval for $Q_p$ is computed using Theorem 2. Also presented in Table 3 are the nonparametric estimate of 10% trimmed mean, its standard error and 95% confidence interval; and parametric and nonparametric $(0.90, 0.95)$ tolerance intervals. The computations involving trimmed mean and tolerance interval are described in Sections 3 and 5, respectively. The parametric tolerance interval is computed using Mee's approach in Krishnamoorthy and Mathew (2009, sec. 4.5). All these confidence and tolerance intervals are also marked on the histogram in Figure 1.

We note that there are substantial differences between the parametric and the nonparametric estimates reported in Table 3. In particular, due to the long right tail of the distribution, it is reasonable that the MLE of $Q_{0.5}$, which is the overall sample mean of the data, is greater than the nonparametric median estimate. Moreover, the nonparametric estimates of $Q_{0.9}$ and $Q_{0.99}$ and the nonparametric tolerance interval are to the right of their parametric counterparts for the same reason. Overall, these findings confirm that the nonparametric estimates are more consistent with the observed data distribution than the normality-based estimates even though the latter lead to smaller standard errors for $\hat{Q}_p$ and a shorter tolerance interval.

Using the nonparametric estimates, we conclude that median of the distribution of systolic blood pressure measurements made by the semi-automatic monitor is 135 (95% confidence interval: $[128, 142]$), its 90th percentile is 192 (95% confidence interval: $[181, 217]$) and its 99th percentile is 228 (95% confidence interval: $[226, 228]$). Further, 90% of the distribution of measurements is contained in $[94, 224]$ with 95% confidence. The 10% trimmed mean of 140 (95% confidence interval: [133, 147]) provides another estimate of the center of the distribution — it is shifted to the right of the median due to right-skewness in the distribution.

Finally a remark is in order about the nonparametric confidence interval for $Q_{0.99}$. This interval is not expected to be accurate as the number of subjects in these data ($n = 85$) is considerably smaller than $n = 250$ needed to achieve satisfactory coverage probability (see Section 6.2). Note also that the upper endpoint of this interval coincides with $\hat{Q}_{0.99}$ and the two equal 228, the largest observation in the data. This is due to the relatively small $n$ and that the interval endpoints need to be observations in the sample (see Theorem 2).

# 8 Technical details and proofs

This section is devoted to proving Theorems 1-4. For the functionals $T$, $T_1$ and $T_2$ given by (8), we can write

$$\begin{aligned} n^{1/2}[T(F_n) - T(F)] &= n^{1/2}[T_1(F_n) - T_1(F)] + n^{1/2}[T_2(F_n) - T_2(F)] \\ &= n^{1/2} \sum_{i=1}^{n} w_i \sum_{j=1}^{k_i} IC(X_{ij}, F, T) + n^{1/2}\Delta_{1n} + n^{1/2}\Delta_{2n}, \end{aligned} \tag{28}$$

where $\Delta_{ln}$ represents the remainder term

$$\Delta_{ln} = T_l(F_n) - T_l(F) - \sum_{i=1}^{n} w_i \sum_{j=1}^{k_i} IC(X_{ij}, F, T_l), \quad l = 1, 2, \tag{29}$$

and the influence curves are given by (10) and (11). The following results hold for the terms on the RHS of (28).

**Proposition 1.** *Let $\sigma^2$ be as defined in* (13)*. Then, under the assumptions A.1 to A.7,*

$$n^{1/2} \sum_{i=1}^{n} w_i \sum_{j=1}^{k_i} IC(X_{ij}, F, T) \xrightarrow{d} N(0, \sigma^2).$$

**Proposition 2.** *Under the assumptions A.1 to A.7, $n^{1/2}\Delta_{1n} = o_p(1)$.*

**Proposition 3.** *Let $G$ be the bivariate c.d.f. of $(\tilde{X}_1, \tilde{X}_2)$. Assume that $G$ is continuous at $(Q_{p_l}, Q_{p_l})$ and $F'(Q_{p_l}) > 0$, for each $l = 1, \ldots, r$. Then, under the assumptions A.1 to A.6, $n^{1/2}\Delta_{2n} = o_p(1)$.*

We prove these results in the next three sections. But first let us use them to quickly establish Theorem 1.

***Proof of Theorem 1.*** The result follows immediately from (28) by applying Propositions 1, 2 and 3, and Slutsky's theorem. □

## 8.1 Proof of Proposition 1

Let $\eta_i = \sum_{j=1}^{k_i} IC(X_{ij}, F, T)$ and $T_{ni} = n^{1/2} w_i \eta_i$, $i = 1, \ldots, n$. These $\eta_i$ are independent with mean zero and variance $\psi^2(k_i)$, defined in (12). The finiteness of this variance is ensured by the second part of assumption A.7 (Shao 2003, exer. 5.34). Note also that $\sum_{i=1}^{n} T_{ni} = n^{1/2} \sum_{i=1}^{n} w_i \sum_{j=1}^{k_i} IC(X_{ij}, F, T)$, and $\sigma_n^2$, given by (13), is the variance of this sum. Next, for the $\delta > 0$ assumed in A.6, we can write:

$$\frac{\sum_{i=1}^{n} E\left|T_{ni}\right|^{2+\delta}}{\sigma_n^{2+\delta}} = \frac{n^{\frac{2+\delta}{2}} \sum_{i=1}^{n} w_i^{2+\delta} E\left|\eta_i\right|^{2+\delta}}{n^{\frac{2+\delta}{2}} \left(\sum_{i=1}^{n} w_i^2 \psi^2(k_i)\right)^{\frac{2+\delta}{2}}} \leq \frac{\max_{1 \leq i \leq n} w_i^{2+\delta} \sum_{i=1}^{n} E\left|\eta_i\right|^{2+\delta}}{\min_{1 \leq i \leq n} w_i^{2+\delta} \left(\sum_{i=1}^{n} \psi^2(k_i)\right)^{\frac{2+\delta}{2}}}. \tag{30}$$

Further, from assumptions A.3 and A.4, we have:

$$\frac{\sum_{i=1}^{n} E\left|\eta_i\right|^{2+\delta}}{\left(\sum_{i=1}^{n} \psi^2(k_i)\right)^{\frac{2+\delta}{2}}} = \frac{n \sum_{i=1}^{n} \frac{E|\eta_i|^{2+\delta}}{n}}{n^{\frac{2+\delta}{2}} \left(\sum_{i=1}^{n} \frac{\psi^2(k_i)}{n}\right)^{\frac{2+\delta}{2}}} \sim n^{-\frac{\delta}{2}} \frac{\sum_{k=1}^{k^*} \mu(k)\, E\left|\eta_k\right|^{2+\delta}}{\left(\sum_{k=1}^{k^*} \mu(k)\, \psi^2(k)\right)^{\frac{2+\delta}{2}}}.$$

The rightmost ratio is free of $n$. From (30) and assumption A.6, this means $\sum_{i=1}^{n} E|T_{ni}|^{2+\delta} = o(\sigma_n^{2+\delta})$. Therefore, from the Liapounov theorem (Shao 2003, p. 69), $\sum_{i=1}^{n} T_{ni}/\sigma_n \xrightarrow{d} N(0, 1)$. The result now holds from Slutsky's theorem since $\sigma^2$ is the limit of $\sigma_n^2$. □

## 8.2 Proof of Proposition 2

In this section, we deduce the desired $n^{1/2}\Delta_{1n} = o_p(1)$ from a more general result, which extends the results of Fernholz (1983, chap. 4) about the remainder term in the von Mises expansion of a *Hadamard differentiable* functional from i.i.d. data to repeated measurements data.

Let $Y_{ij} = F(X_{ij})$, so that the $Y_{ij}$ are distributed uniformly on $[0, 1]$. Also, let $U$ be the c.d.f. of this uniform distribution. Define the counterpart of $F_n$ for the $Y_{ij}$ as:

$$U_n(x) = \sum_{i=1}^{n} w_i \sum_{j=1}^{k_i} I(Y_{ij} \le x). \tag{31}$$

Next, let $\mathbb{D}[0, 1]$ be the space of *cadlag* functions (i.e., right continuous functions with left-hand limits) on $[0, 1]$. We assume that $\mathbb{D}$ is equipped with the sup norm $\|\cdot\|_\infty$. Suppose we have a functional $\tau : \mathbb{D}[0, 1] \longmapsto \mathbb{R}$ that is Hadamard differentiable at $U \in \mathbb{D}[0, 1]$ with derivative $\tau'_U$. From the von Mises expansion, we have:

$$n^{1/2}[\tau(U_n) - \tau(U)] = n^{1/2}\tau'_U(U_n - U) + n^{1/2}\,\mathrm{Rem}(U_n - U). \tag{32}$$

The remainder term converges in probability to zero from the following result.

**Proposition 4.** *Suppose the assumptions A.1 to A.5 hold. Then, for the remainder term in the von Mises expansion* (32) *of a Hadamard differentiable functional* $\tau$*, we have:* $n^{1/2}\,Rem(U_n - U) = n^{1/2}\Delta_{1n} = o_p(1)$.

Before proving this result, let us first use it to establish Proposition 2.

***Proof of Proposition 2.*** Since $F_n = U_n \circ F$ and $F = U \circ F$, the statistical functional $T_1$, defined in (8), induces a functional $\mathbb{D}[0, 1] \longmapsto \mathbb{R}$. Take $\tau$ to be this functional, i.e.,

$$T_1(F) = T_1(U \circ F) =: \tau(U), \quad T_1(F_n) = T_1(U_n \circ F) =: \tau(U_n).$$

This $\tau$ is known to be Hadamard differentiable at $U \in \mathbb{D}[0, 1]$ due the first part of assumption A.7 (Fernholz 1983, prop. 7.2.1). Therefore, from the von Mises expansion,

$$n^{1/2}[T_1(F_n) - T_1(F)] = n^{1/2}[\tau(U_n) - \tau(U)] = n^{1/2}\tau'_U(U_n - U) + n^{1/2}\,\mathrm{Rem}(U_n - U). \tag{33}$$

Since $U_n = F_n \circ F^{-1}$, $U = F \circ F^{-1}$, and $\tau'_U$ is linear by definition, we can write:

$$\tau'_U(U_n - U) = \tau'_U[(F_n - F) \circ F^{-1}] = \sum_{i=1}^{n} w_i \sum_{j=1}^{k_i} \tau'_U[(\delta_{X_{ij}} - F) \circ F^{-1}] = \sum_{i=1}^{n} w_i \sum_{j=1}^{k_i} IC(X_{ij}, F, T_1),$$

where the last equality follows from Fernholz (1983, lem. 4.4.1). Next, a comparison of (29) and (33) shows that $n^{1/2}\Delta_{1n} = n^{1/2}\,\mathrm{Rem}(U_n - U)$. The result now follows from Proposition 4. □

To prove Proposition 4, we begin by establishing convergence of the weighted empirical process $U_n$. Let $\mathbb{G}$ be a continuous Gaussian process with mean zero and covariance $\sum_{k=1}^{k^*}\mu(k)\theta^2(k)\varphi_k^2(x,y)$. Here $k^*$, $\mu(k)$ and $\theta(k)$ are as defined in assumptions A.3-A.5, and $\varphi_k^2(x,\,y) = \mathrm{cov}[\sum_{l=1}^{k} I(F(\tilde{X}_l) \le x), \sum_{l=1}^{k} I(F(\tilde{X}_l) \le y)]$. The following result is proven in Assaad (2012) by essentially proceeding along the lines of Olsson and Rootzen (1996, thm. 3.1).

**Lemma 1.** *Suppose that the assumptions A.1-A.5 hold. Then, for $U_n$ defined in (31), we have: $n^{1/2}(U_n - U) \xrightarrow{d} \mathbb{G}$ in $\mathbb{D}[0,1]$.*

Next, it is well-known that $U_n$ is not a random element of $\mathbb{D}[0,1]$ as this space when equipped with $\|\cdot\|_\infty$ norm is complete but not separable (Fernholz 1983, chap. 4). We deal with this difficulty as in Fernholz by studying a continuous version $U_n^*$ of $U_n$. Let $Y_{(0)} = 0$, $Y_{(1)} = F(X_{(1)}), \ldots, Y_{(N)} = F(X_{(N)})$, $Y_{(N+1)} = 1$. The intervals $[Y_{(i-1)}, Y_{(i)}]$, $i = 1, \ldots, N+1$, form a partition of $[0,1]$. Next, let $p_{i-1}$ be an arbitrary probability mass that is less than the weight of $X_{(i)}$, $i = 1, \ldots, N$, and take $p_N = 1 - (p_0 + \ldots + p_{N-1})$ so that $\sum_{i=0}^{N} p_i = 1$. Define

$$U_n^*(x) = \left(\bar{p}_{i-2} + p_{i-1}\frac{x - Y_{(i-1)}}{Y_{(i)} - Y_{(i-1)}}\right) I_{[Y_{(i-1)},Y_{(i)}]}(x), \tag{34}$$

with $\bar{p}_j = \sum_{i=0}^{j} p_i$. This $U_n^*$ is continuous since $Y_{(i)} \ne Y_{(j)}$ for $i \ne j$ (with probability 1). Essentially this $U_n^*$ distributes the probability mass $p_{i-1}$ uniformly in interval $i$ for each $i$. The way $p_{i-1}$ are defined ensures:

$$\|U_n^* - U_n\|_\infty \le \max_{i=1,\ldots,n} w_i = \max_{k=1,\ldots,k^*} w(k) \quad (a.s.). \tag{35}$$

Let $\mathbb{C}[0,1]$ denote the space of continuous functions on $[0,1]$ equipped with the sup-norm $\|\cdot\|_\infty$. Since this space is complete and separable, Billingsley (1968, p. 84) and (34) imply that $U_n^*$ is a random element of $\mathbb{C}[0,1]$. In addition, from (35) and the assumption A.5, it can be seen that:

$$n^{1/2}\|U_n - U_n^*\|_\infty = o_{p_*}(1). \tag{36}$$

Here we use the inner probability measure $P_*$ corresponding to $P$ instead of $P$ as $U_n$ and hence $n^{1/2}(U_n - U_n^*)$ is not a random element of $\mathbb{D}[0,1]$. We can now state the following results.

**Lemma 2.** *The random element $n^{1/2}(U_n^* - U)$ is tight in $\mathbb{C}[0,1]$.*

*Proof.* The fact that $n^{1/2}(U_n - U) \xrightarrow{d} \mathbb{G}$ in $\mathbb{D}[0,1]$ (by Lemma 1) implies from (36) and van der Vaart (1998, thm 18.10 (iv)) that $n^{1/2}(U_n^* - U) \xrightarrow{d} \mathbb{G}$ in $\mathbb{C}[0,1]$. Therefore, $n^{1/2}(U_n^* - U)$ is relatively compact. Now the tightness follows from Prohorov's theorem as $\mathbb{C}[0,1]$ is separable and complete. $\square$

**Lemma 3.** *$\forall\, \epsilon > 0$, $\exists$ a compact set $K \subset \mathbb{D}[0,1]$, $M > 0$ and $n_0 \in \mathbb{N}$ such that $\forall\, n \geq n_0$, we have:*

$$P_*\{d_\infty(\sqrt{n}(U_n - U),\, K) \leq M/n^{1/2}\} > 1 - \epsilon,$$

*where $d_\infty(H, K) = \inf_{E \in K} \|H - E\|_\infty$ for $H \in \mathbb{D}[0,1]$ and $K \subset \mathbb{D}[0,1]$.*

*Proof.* From (35) and A.5, $\exists\, M > 0$ and $n_0 \in \mathbb{N}$ such that $\|U_n - U_n^*\|_\infty < M/n$, almost everywhere $\forall\, n \geq n_0$. Further, by Lemma 2, $\exists$ a compact set $K \subset \mathbb{C}[0,1]$ such that $\forall\, n$:

$$P_*[n^{1/2}(U_n^* - U) \in K] > 1 - \epsilon. \tag{37}$$

This $K$ is also compact in $\mathbb{D}[0,1]$ as $\mathbb{C}[0,1] \subset \mathbb{D}[0,1]$. Now, define the events $A = \{n^{1/2}(U_n^* - U) \in K\}$, $B = \{\|U_n^* - U_n\|_\infty < M/n\}$ and $C = \{d_\infty(n^{1/2}[U_n - U],\, K) \leq M/n^{1/2}\}$. The event $A \cap B$ is a subset of the event $C$ because if $A$ and $B$ occur, then

$$d_\infty(n^{1/2}[U_n - U],\, K) \leq d_\infty(n^{1/2}[U_n - U], n^{1/2}[U_n^* - U]) = n^{1/2}\|U_n^* - U_n\|_\infty < M/n^{1/2}.$$

The result now follows from (37) by noticing that $P_*(A \cap B) = P_*(A)$ for all $n \geq n_0$. $\square$

Next, we state a result of Fernholz (1983) after making minor modifications to it to suit our purpose.

**Lemma 4.** *[Fernholz (1983, lem. 4.3.1)] Let $Q : \mathbb{D}[0,1] \times \mathbb{R} \longmapsto \mathbb{R}$ and suppose that for any compact set $K \subset \mathbb{D}[0,1]$, $\lim_{t \to 0} Q(H, t) = 0$ uniformly for $H \in K$. Let $\epsilon > 0$, and let $\delta_n$ be a sequence of numbers with $\delta_n \downarrow 0$. Then for any compact set $K \subset \mathbb{D}[0,1]$, $\exists\, n_0$ such that $\forall\, n \geq n_0$, if $d_\infty(H, K) \leq \delta_n$ then $|Q(H, r\delta_n)| < \epsilon$, for any constant $r \in \mathbb{R}$.*

We are now ready to prove Proposition 4.

***Proof of Proposition 4.*** Let $\epsilon > 0$, and $C_n$ be the event $\{d_\infty(n^{1/2}[U_n - U], K) \leq M/n^{1/2}\}$. From Lemma 3, we can choose a compact set $K \subset \mathbb{D}[0,1]$ and $M > 0$ such that $P_*(C_n) > 1 - \epsilon/2$, $\forall\, n \geq n_0$. Further, since $P_*$ is an inner probability measure, we can find measurable sets $E_n$ such that $E_n \subset C_n$ and $P(E_n) > P_*(C_n) - \epsilon/2$. Thus, we have:

$$P(E_n) > P_*(C_n) - \epsilon/2 > 1 - \epsilon, \quad \forall\, n \geq n_0. \tag{38}$$

Next, let $\text{Rem}(H) = \tau(U+H) - \tau(U) - \tau'_U(H)$. The Hadamard differentiability of $\tau$ at $U$ implies that $\text{Rem}(tH)/t \to 0$ as $t \to 0$, uniformly for $H \in K$ found earlier. Now, upon applying Lemma 4 with

$$Q(H,t) = \text{Rem}(tH)/t, \delta_n = M/n^{1/2}, r = 1/M \text{ and } H = n^{1/2}(U_n - U),$$

we can find $n_1$ such that $\forall\, n > n_1$, $d_\infty(n^{1/2}[U_n - U], K) \leq M/n^{1/2}$ implies $|Q(n^{1/2}[U_n - U], 1/n^{1/2})| < \epsilon$. Therefore, for $n > n_2 = \max\{n_0, n_1\}$, we have:

$$\begin{aligned} P_*\{|Q(n^{1/2}[U_n - U], 1/n^{1/2})| < \epsilon\} &= P_*\{n^{1/2}|\text{Rem}(U_n - U)| < \epsilon\} \\ &= P\{n^{1/2}|\text{Rem}(U_n - U)| < \epsilon\} \geq P_*(C_n) > P(E_n) > 1 - \epsilon, \end{aligned}$$

where the second equality is due to the fact that $\text{Rem}(U_n - U)$ is a random element of $\mathbb{D}[0,1]$ even though $U_n$ is not (see Fernholz 1983, p. 40), and the last inequality is from (38). Hence, $n^{1/2}\text{Rem}(U_n - U) \xrightarrow{p} 0$. □

## 8.3 Proof of Proposition 3

As seen next, the result in Proposition 3 readily follows from the Bahadur representation in Theorem 3.

***Proof of Proposition 3.*** For $l = 1, \ldots, r$, define:

$$\Delta_{2n,l} = \hat{Q}_{p_l} - Q_{p_l} - \{p_l - F_n(Q_{p_l})\}/f(Q_{p_l}) = \hat{Q}_{p_l} - Q_{p_l} - \sum_{i=1}^{n} w_i \sum_{j=1}^{k_i} IC(X_{ij}, F, Q_{p_l}),$$

where the last equality follows from (16). Using (29), we can write, $\Delta_{2n} = \sum_{l=1}^{r} a_l \Delta_{2n,l}$. Next, for each $l$, taking the constant sequence $p^{(n)} = p_l \;\forall\, n$ in Theorem 3 yields $n^{1/2}\Delta_{2n,l} = o_p(1)$. This implies $n^{1/2}\Delta_{2n} = o_p(1)$ and hence the result holds. □

## 8.4 Proof of Theorem 2

We first present two results that are needed for proving Theorem 2.

**Lemma 5.** *Under the assumptions of Theorem 2, $\hat{r}_n^2 \xrightarrow{p} \sigma^2 f^2(Q_p)$ as $n \to \infty$, where $\sigma^2$ is given by (17).*

*Proof.* It suffices to show that $\hat{\rho}(\hat{Q}_p, \hat{Q}_p) \xrightarrow{p} \rho(Q_p, Q_p)$, defined by (4), since then

$$\hat{r}_n^2 = p(1-p) \sum_{k=1}^{k^*} k\{1 + (k-1)\hat{\rho}(\hat{Q}_p, \hat{Q}_p)\}\mu_n(k)\{nw(k)\}^2 \xrightarrow{p} \sigma^2 f^2(Q_p).$$

Under the assumptions, $\rho(Q_p, Q_p)$ is continuous at $(Q_p, Q_p)$. In addition, as $n^{1/2}(\hat{Q}_p - Q_p) \xrightarrow{d} N(0, \sigma^2)$ from Theorem 1, we have $\hat{Q}_p = O_p(1)$, implying that for a given $\epsilon > 0$, $\exists\, M_\epsilon > 0$ such that

$$\lim_{n\to\infty} P(|\hat{Q}_p| \le M_\epsilon) \to 1. \tag{39}$$

To prove the convergence of $\hat{\rho}$, note that

$$\big|\hat{\rho}(\hat{Q}_p, \hat{Q}_p) - \rho(Q_p, Q_p)\big| \le \big|\hat{\rho}(\hat{Q}_p, \hat{Q}_p) - \rho(\hat{Q}_p, \hat{Q}_p)\big| + \big|\rho(\hat{Q}_p, \hat{Q}_p) - \rho(Q_p, Q_p)\big|.$$

The second term on the right goes to zero in probability due to the continuity of $\rho$. Thus it just remains to show that the first term also goes to zero in probability. To see this, we have for $\epsilon > 0$,

$$\begin{aligned} P(\big|\hat{\rho}(\hat{Q}_p, \hat{Q}_p) - \rho(\hat{Q}_p, \hat{Q}_p)\big| > \epsilon) &= P(\big|\hat{\rho}(\hat{Q}_p, \hat{Q}_p) - \rho(\hat{Q}_p, \hat{Q}_p)\big| > \epsilon,\, |\hat{Q}_p| \le M_\epsilon) \\ &\qquad + P(\big|\hat{\rho}(\hat{Q}_p, \hat{Q}_p) - \rho(\hat{Q}_p, \hat{Q}_p)\big| > \epsilon,\, |\hat{Q}_p| > M_\epsilon) \\ &\le P\big(\sup_{|x|\le M_\epsilon} \big|\hat{\rho}(x,x) - \rho(x,x)\big| > \epsilon\big) + P(|\hat{Q}_p| > M_\epsilon). \end{aligned}$$

The first term on the right goes to zero from Olsson and Rootzen (1996, p. 1563). The second term goes to zero from (39). This establishes the result. □

**Lemma 6.** *Suppose the assumptions of Theorem 2 hold.*

*(a) Let $p^{(n)}$ be a sequence of probabilities such that $p^{(n)} = p + c/n^{1/2} + o(1/n^{1/2})$. Then as $n \to \infty$, $n^{1/2}(\hat{Q}_{p^{(n)}} - \hat{Q}_p) \xrightarrow{p} c/f(Q_p)$.*

*(b) Let $\hat{p}^{(n)}$ be a sequence of probabilities such that $\hat{p}^{(n)} = p + \hat{c}_n/n^{1/2}$, where $\hat{c}_n \xrightarrow{p} c$. Then as $n \to \infty$, $n^{1/2}(\hat{Q}_{\hat{p}^{(n)}} - \hat{Q}_p) \xrightarrow{p} c/f(Q_p)$.*

*Proof.* The part (a) can be proved by adapting van der Vaart (1998, lem. 21.7) to deal with repeated measurements (see Assaad 2012). Here we focus on using (a) to prove (b). Fix $\epsilon > 0$ and consider,

$$\begin{aligned} P\big(n^{1/2}|\hat{p}^{(n)} - p^{(n)}| \le \epsilon\big) &= P\big(p^{(n)} - \epsilon/n^{1/2} \le \hat{p}^{(n)} \le p^{(n)} + \epsilon/n^{1/2}\big) \\ &\le P\big(\hat{Q}_{p^{(n)}-\epsilon/n^{1/2}} \le \hat{Q}_{\hat{p}^{(n)}} \le \hat{Q}_{p^{(n)}+\epsilon/n^{1/2}}\big). \end{aligned}$$

The probabilities above go to one since $n^{1/2}\big(\hat{p}^{(n)} - p^{(n)}\big) = \hat{c}_n - c + o(1) \xrightarrow{p} 0$. As a result,

$$\lim_{n\to\infty} P\big\{n^{1/2}(\hat{Q}_{p^{(n)}-\epsilon/n^{1/2}} - \hat{Q}_p) \le n^{1/2}(\hat{Q}_{\hat{p}^{(n)}} - \hat{Q}_p) \le n^{1/2}(\hat{Q}_{p^{(n)}+\epsilon/n^{1/2}} - \hat{Q}_p)\big\} = 1. \tag{40}$$

Next, we can deduce from (a) that $n^{1/2}(\hat{Q}_{p^{(n)}-\epsilon/n^{1/2}} - \hat{Q}_p) \xrightarrow{p} (c-\epsilon)/f(Q_p)$ and $n^{1/2}(\hat{Q}_{p^{(n)}+\epsilon/n^{1/2}} - \hat{Q}_p) \xrightarrow{p} (c+\epsilon)/f(Q_p)$. Therefore,

$$\lim_{n\to\infty} P\big(n^{1/2}(\hat{Q}_{p^{(n)}-\epsilon/n^{1/2}} - \hat{Q}_p) - (c-\epsilon)/f(Q_p) \geq -\epsilon\big) = 1, \tag{41}$$

$$\lim_{n\to\infty} P\big(n^{1/2}(\hat{Q}_{p^{(n)}+\epsilon/n^{1/2}} - \hat{Q}_p) - (c+\epsilon)/f(Q_p) \leq \epsilon\big) = 1. \tag{42}$$

Let $A_n$, $B_n$ and $C_n$ denote the events in (40), (41) and (42), respectively. Notice that the event $A_n \cap B_n \cap C_n$ implies the event

$$-\epsilon\{1 + 1/f(Q_p)\} \leq n^{1/2}(\hat{Q}_{\hat{p}^{(n)}} - \hat{Q}_p) - c/f(Q_p) \leq \epsilon\{1 + 1/f(Q_p)\}.$$

From Lehmann (1999, lem. 2.1.2), its probability goes to one since each of the three probabilities, $P(A_n)$, $P(B_n)$ and $P(C_n)$, goes to one. This establishes the result as $\epsilon > 0$ is arbitrary. □

We are now ready to prove Theorem 2.

***Proof of Theorem 2.*** We can write the coverage probability as

$$P(\hat{Q}_{\hat{l}_n} \leq Q_p \leq \hat{Q}_{\hat{u}_n}) = P\{n^{1/2}(\hat{Q}_p - \hat{Q}_{\hat{u}_n}) \leq n^{1/2}(\hat{Q}_p - Q_p) \leq n^{1/2}(\hat{Q}_p - \hat{Q}_{\hat{l}_n})\}.$$

From Theorem 1, we know that $n^{1/2}(\hat{Q}_p - Q_p) \xrightarrow{d} N(0,\sigma^2)$. Therefore, it suffices to show that $n^{1/2}(\hat{Q}_{\hat{u}_n} - \hat{Q}_p) \xrightarrow{p} z_{1-\beta/2}\,\sigma$ and $n^{1/2}(\hat{Q}_{\hat{l}_n} - \hat{Q}_p) \xrightarrow{p} -z_{1-\beta/2}\,\sigma$ as then the result follows from Slutsky's theorem. To get the limits of the $\hat{Q}$ differences, take $\hat{c}_n = z_{1-\beta/2}\hat{r}_n$ so that $\hat{l}_n = p - \hat{c}_n/n^{1/2}$ and $\hat{u}_n = p + \hat{c}_n/n^{1/2}$. Next, an application of Lemma 5 gives $\hat{c}_n \xrightarrow{p} z_{1-\beta/2}\,\sigma f(Q_p)$. The desired result now follows from part (b) of Lemma 6 upon taking $\hat{p}^{(n)} = \hat{l}_n$ and $\hat{p}^{(n)} = \hat{u}_n$. □

## 8.5 Proof of Theorem 3

The following lemma is needed to prove Theorem 3.

**Lemma 7.** *[Ghosh, 1971] Let $\{V_n\}$ and $\{W_n\}$ be two sequences of random variables satisfying the following conditions:*

$$W_n = O_p(1);\ \textit{and}\ \forall\, t \text{ and } \forall\, \epsilon > 0,\ \lim_{n\to\infty} P(V_n \leq t,\, W_n \geq t+\epsilon) = 0,\ \lim_{n\to\infty} P(W_n \leq t,\, V_n \geq t+\epsilon) = 0. \tag{43}$$

*Then $V_n - W_n = o_p(1)$.*

***Proof of Theorem 3.*** We proceed along the lines of Ghosh (1971) to get this result. Let $\gamma_n = Q_p + (p^{(n)} - p)/f(Q_p)$, $V_n = n^{1/2}(\hat{Q}_{p^{(n)}} - \gamma_n)$ and $W_n = n^{1/2}\{p - F_n(Q_p)\}/f(Q_p)$. Since

$$V_n - W_n = n^{1/2}(\hat{Q}_{p^{(n)}} - Q_p) - n^{1/2}\{p^{(n)} - F_n(Q_p)\}/f(Q_p),$$

it is enough to verify that $V_n$ and $W_n$ satisfy (43) as then the result is an immediate consequence of Lemma 7.

From (16), we can write $W_n = n^{1/2}\sum_{i=1}^{n} w_i \sum_{j=1}^{k_i} IC(X_{ij}, F, Q_p)$. This $W_n$ can be shown to be asymptotically normal by proceeding as in Proposition 1. Thus, $W_n = O_p(1)$. Next, for a given $t$, let

$$Z_{t,n} = n^{1/2}\big\{F(\gamma_n + t/n^{1/2}) - F_n(\gamma_n + t/n^{1/2})\big\}/f(Q_p), \quad t_n = n^{1/2}\big\{F(\gamma_n + t/n^{1/2}) - p^{(n)}\big\}/f(Q_p).$$

It can be seen that the event $\{V_n \le t\} \subset \{Z_{t,n} \le t_n\}$, and $\lim_{n\to\infty} t_n = t$ as $n^{1/2}(p^{(n)} - p) = O(1)$. Moreover, the random variable $Z_{t,n} - W_n$ has mean zero and variance

$$E[Z_{t,n} - W_n]^2 = \frac{n}{f^2(Q_p)}\mathrm{var}\big[F_n(Q_p) - F_n(\gamma_n + t/n^{1/2})\big] = \frac{n}{f^2(Q_p)}\sum_{i=1}^{n} w_i^2 J_n(k_i), \tag{44}$$

where $J_n(k_i) = \mathrm{var}\big[\sum_{j=1}^{k_i}\{\delta_{X_{ij}}(Q_p) - \delta_{X_{ij}}(\gamma_n + t/n^{1/2})\}\big]$, which, from A.2, can be written as

$$k_i \mathrm{var}\big[\delta_{\tilde{X}_1}(Q_p) - \delta_{\tilde{X}_1}(\gamma_n + t/n^{1/2})\big] + 2\binom{k_i}{2}\mathrm{cov}\big[\delta_{\tilde{X}_1}(Q_p) - \delta_{\tilde{X}_1}(\gamma_n + t/n^{1/2}), \delta_{\tilde{X}_2}(Q_p) - \delta_{\tilde{X}_2}(\gamma_n + t/n^{1/2})\big].$$

Upon simplifying it using the facts that $\mathrm{var}[\delta_{\tilde{X}_1}(a)] = F(a)\{1 - F(a)\}$ and $\mathrm{cov}[\delta_{\tilde{X}_1}(a), \delta_{\tilde{X}_2}(b)] = G(a,b) - F(a)F(b)$, and applying continuity of $G$ at $(Q_p, Q_p)$, we get $\lim_{n\to\infty} J_n(k_i) = 0$. Next, by writing (44) as

$$E[Z_{t,n} - W_n]^2 = \frac{1}{f^2(Q_p)}\sum_{k=1}^{k^*}\{nw(k)\}^2\mu_n(k)J_n(k),$$

it follows from A.4 and A.5 that $\lim_{n\to\infty} E[Z_{t,n} - W_n]^2 = 0$. Therefore, $Z_{t,n} - W_n = o_p(1)$. This together with $t_n \to t$ imply that $P(Z_{t,n} \le t_n,\, W_n \ge t + \epsilon) \to 0$, where $\epsilon > 0$. Further, since $\{V_n \le t\} \subset \{Z_{t,n} \le t_n\}$, we can deduce that $P(V_n \le t,\, W_n \ge t + \epsilon) \to 0$, $\forall\, t$, $\epsilon > 0$. A similar argument shows that $P(W_n \le t,\, V_n \ge t + \epsilon) \to 0$, $\forall\, t, \epsilon > 0$. Thus, $V_n$ and $W_n$ satisfy conditions (43) of Lemma 7, which completes the proof. □

## 8.6 Proof of Theorem 4

To prove (a), take $r = 2$ in the general $L$-statistic formula (7) and set the continuous part $T_1$ equal to zero to get $T(F_n) = a_1\hat{Q}_{p_1} + a_2\hat{Q}_{p_2}$, $a_1, a_2 \in \mathbb{R}$, $(a_1, a_2) \ne (0,0)$. In this case, $T(F) = a_1 Q_{p_1} + a_2 Q_{p_2}$. From

Theorem 1, $n^{1/2}[T(F_n) - T(F)] \xrightarrow{d} N(0, \sigma^2)$, where $\sigma^2$, obtained using (11), (12) and (13), can be written as

$$\sigma^2 = a_1^2 \frac{\nu_1^2(p_1)}{f^2(Q_{p_1})} + 2a_1a_2 \frac{\nu_{12}(p_1, p_2)}{f(Q_{p_1})f(Q_{p_2})} + a_2^2 \frac{\nu_2^2(p_2)}{f^2(Q_{p_2})},$$

with $\nu_1$, $\nu_2$ and $\nu_{12}$ as defined in (22). Since this result holds for any $(a_1, a_2) \neq (0, 0)$, we can deduce from the Cramer-Wold device (van der Vaart 1998, p. 16) that $n^{1/2}(\hat{Q}_{p_1} - Q_{p_1}, \hat{Q}_{p_2} - Q_{p_2})$ jointly converges in distribution to a bivariate normal distribution with mean $(0, 0)$, variance $(\nu_1^2(p_1)/f^2(Q_{p_1}), \nu_2^2(p_2)/f^2(Q_{p_2}))$ and covariance $\nu_{12}(p_1, p_2)/\{f(Q_{p_1})f(Q_{p_2})\}$. Next, take $h(Q_{p_1}, Q_{p_2}) = F(Q_{p_2}) - F(Q_{p_1})$ so that $n^{1/2}[F(\hat{Q}_{p_2}) - F(\hat{Q}_{p_1}) - (p_2 - p_1)] = n^{1/2}[h(\hat{Q}_{p_1}, \hat{Q}_{p_2}) - h(Q_{p1}, Q_{p2})]$. From the bivariate delta method (Lehmann 1999, p. 295), this quantity converges in distribution to $N(0, \nu^2(p_1, p_2))$, completing the proof.

To prove (b), note that from Theorem 1, $n^{1/2}(\hat{Q}_{p_l} - Q_{p_l}) \xrightarrow{d} N(0, \nu_l^2(p_l)/f^2(Q_{p_l}))$, $l = 1, 2$. It now follows from the usual delta method that $n^{1/2}(F(\hat{Q}_{p_l}) - p_l) = n^{1/2}(F(\hat{Q}_{p_l}) - F(Q_{p_l})) \xrightarrow{d} N(0, \nu_l^2(p_l))$, $l = 1, 2$. □


## Acknowledgement

The authors would like to thank the reviewers and the Associate Editor for providing thoughtful comments on this work. They have led to substantial improvements in this article.


## References


[1] Assaad, H. (2012). $L$-statistics for repeated measurements data and their applications. Ph.D. dissertation, University of Texas at Dallas.

[2] Azzalini, A. (1985). A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics*, **12**, 171-178.

[3] Billingsley, P. (1968). *Convergence of Probability Measures*. John Wiley, New York.

[4] Bland, J. M. and Altman, D. G. (1999). Measuring agreement in method comparison studies. *Statistical Methods in Medical Research*, **8**, 135-160.

[5] Burdick, R., Borror, C. and Montgomery, D. (2005). *Design and Analysis of Gauge R&R Studies: Making Decisions with Confidence Intervals in Random and Mixed Anova Models*, ASA-SIAM Series on Statistics & Applied Probability, SIAM, Philadelphia, ASA, Alexandria, VA.

[6] David, H. A. and Nagaraja, H. N. (2003). *Order Statistics*, 3rd edition. John Wiley, New York.

[7] Dunn, G. (1989). *Design and Analysis of Reliabilities studies: The Statistical Evaluation of Measurement Errors*. Oxford University Press, New York.

[8] Fernholz, L. (1983). *von Mises Calculus for Statistical Functionals*. Springer, New York.

[9] Fleiss, J. L. (1986). *The Design and Analysis of Clinical Experiments*. John Wiley, New York.

[10] Ghosh, J. K. (1971). A new proof of the Bahadur representation of quantiles and an application. *Annals of Mathematical Statistics*, **42**, 1957-1961.

[11] Guttman, I. (1988). Statistical tolerance regions. In *Encyclopedia of Statistical Sciences*, **9**, pp. 272-287, Kotz, S., Johnson, N. L. and Read, C. B. (Editors), John Wiley, New York

[12] Huang, M. L. and Brill, P. H. (1999). A level crossing quantile estimation method. *Statistics & Probability Letters*, **45**, 111-119.

[13] Huber, P. J. (1981). *Robust Statistics*. John Wiley, New York.

[14] Hutson, A. (2003). Nonparametric estimation of normal ranges given one-way ANOVA random effects assumptions. *Statistics & Probability Letters*, **64**, 415-424.

[15] Karlin, S., Cameron, E . C. and Williams, P. T. (1981). Sibling and parent-offspring correlation estimation with a variable family size, *Proceedings of the National Academy of Sciences*, **78**, 2664-2668.

[16] Keselman, H. J., Kowalchuk, R. K., Algina, J., Lix, L. M. and Wilcox, R. R. (2000). Testing treatment effects in repeated measures designs: Trimmed means and bootstrapping, *British Journal of Mathematical & Statistical Psychology*, **53**, 175-91.

[17] Krishnamoorthy, K. and Mathew, T. (2009). *Statistical Tolerance Regions: Theory, Applications, and Computation*. John Wiley, New York.

[18] Lehmann, E. L. (1999). *Elements of Large-Sample Theory*. Springer, New York.

[19] Olsson, J. and Rootzen, H. (1996). Quantile estimation from repeated measurements. *Journal of the American Statistical Association*, **91**, 1560-1565.

[20] Pinheiro, J. C. and Bates, D. (2000). *Mixed-Effects Models in S and S-PLUS*. Springer, New York.

[21] Pinheiro, J. C., Bates, D., DebRoy, S., Sarkar, D. and the R Development Core Team (2011). nlme: Linear and Nonlinear Mixed Effects Models. R package version 3.1-101.

[22] R Development Core Team (2011). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

[23] Serfling, R. J. (1980). *Approximation Theorems of Mathematical Statistics*. John Wiley, New York.

[24] Shao, J. (2003). *Mathematical Statistics*, 2nd edition. Springer, New York.

[25] Sharma, G. and Mathew, T. (2012). One-sided and two-sided tolerance intervals in general mixed and random effects models using small sample asymptotics. *Journal of the American Statistical Association*, to appear, DOI: 10.1080/01621459.2011.640592

[26] Silverman, B. W. (1986). *Density Estimation for Statistics and Data Analysis*. Chapman & Hall/CRC, Boca Raton.

[27] van der Vaart, A. W. (1998). *Asymptotic Statistics*. Cambridge University Press, New York.

[28] Vardeman, S. B. (1992). What about the other intervals? *The American Statistician*, **46**, 193-197.

[29] Wilcox, R. R. (1994). A one-way random effects model for trimmed means. *Psychometrika*, **59**, 289-306.

[30] Wilcox, R. R. (2012). *Introduction to Robust Estimation and Hypothesis Testing*, 3rd edition. Academic Press, San Diego, CA.

[31] Wilcox, R. R., Keselman, H. J., Muska, J. and Cribbie, R. (2000). Repeated measures ANOVA: Some new results on comparing trimmed means and means, *British Journal of Mathematical & Statistical Psychology*, **53**, 69-82.

| Models for $b_i$, $\epsilon_{ij}$ | $\alpha$ 0.05 $\hat{\xi}_1$ | $\hat{\xi}_2$ | 0.10 $\hat{\xi}_1$ | $\hat{\xi}_2$ | 0.125 $\hat{\xi}_1$ | $\hat{\xi}_2$ |
|---|---|---|---|---|---|---|
| $N, N$ | 0.98 | 0.84 | 0.96 | 0.82 | 0.95 | 0.81 |
| $t_5, N$ | 1.20 | 1.00 | 1.24 | 1.03 | 1.25 | 1.04 |
| $t_3, N$ | 1.65 | 1.41 | 1.78 | 1.54 | 1.82 | 1.57 |
| $N, t_3$ | 1.01 | 0.87 | 0.99 | 0.86 | 0.97 | 0.85 |
| $t_5, t_3$ | 1.18 | 1.04 | 1.22 | 1.08 | 1.23 | 1.09 |
| $t_3, t_3$ | 1.62 | 1.36 | 1.77 | 1.49 | 1.80 | 1.52 |
| $N, t_5$ | 0.98 | 0.83 | 0.95 | 0.81 | 0.95 | 0.80 |
| $t_5, t_5$ | 1.16 | 1.00 | 1.18 | 1.03 | 1.18 | 1.03 |
| $t_3, t_5$ | 1.65 | 1.39 | 1.80 | 1.52 | 1.85 | 1.56 |
| $t_{30}, t_{30}$ | 1.00 | 0.84 | 0.98 | 0.83 | 0.96 | 0.82 |

Table 1: Estimated ARE of the trimmed mean estimate $\hat{\xi}_l$, MSE($\hat{\xi}_{mle}$)/MSE($\hat{\xi}_l$), with respect to $\hat{\xi}_{mle}$, $l = 1, 2$. The "$N$" under models refers to the $N(0,1)$ distribution.

| Models for $b_i, \epsilon_{ij}$ | $p = 0.8$ | | $p = 0.9$ | |
|---|---|---|---|---|
| | $w_{i,1}$ | $w_{i,2}$ | $w_{i,1}$ | $w_{i,2}$ |
| $N, N$ | 94.3 | 93.4 | 94.2 | 92.2 |
| $SN, N$ | 93.3 | 93.3 | 93.1 | 91.4 |
| $t_3, N$ | 94.1 | 94.0 | 92.8 | 92.0 |
| $N, t_3$ | 94.0 | 93.0 | 94.1 | 92.3 |
| $SN, t_3$ | 93.8 | 92.2 | 93.6 | 92.4 |
| $t_3, t_3$ | 93.6 | 93.4 | 92.9 | 92.1 |
| $N, SN$ | 93.9 | 93.5 | 91.7 | 91.8 |
| $SN, SN$ | 92.7 | 92.9 | 93.0 | 91.3 |
| $t_3, SN$ | 93.8 | 93.7 | 92.6 | 93.1 |

Table 2: Proportion of times (in %) the probability content of an asymptotic $(p, 0.95)$ tolerance interval exceeds $p$ in case of an unbalanced design with $n = 60$. The weight functions $w_{i,1}$ and $w_{i,2}$ are given by (2). The "$N$" and "$SN$" under models refer to $N(0, 1)$ and $SN(0, 1, 5)$ distributions, respectively.

| | likelihood-based | | | nonparametric | | |
|---|---|---|---|---|---|---|
| | estimate | S.E. | 95% C.I. | estimate | S.E. | 95% C.I. |
| $Q_{0.5}$ | 143 | 3.4 | $(136, 150)$ | 135 | 3.5 | $(128, 142)$ |
| $Q_{0.9}$ | 185 | 4.6 | $(176, 194)$ | 192 | 7.0 | $(181, 217)$ |
| $Q_{0.99}$ | 219 | 6.5 | $(206, 231)$ | 228 | 3.7 | $(226, 228)$ |
| 10% trimmed mean | - | - | - | 140 | 3.5 | $(133, 147)$ |
| | likelihood-based | | | nonparametric | | |
| $(0.90, 0.95)$ tolerance interval | $(81, 205)$ | | | $(94, 224)$ | | |

Table 3: Comparison of likelihood-based and nonparametric inferences for the blood pressure data. Here "S.E." means "standard error" and "C.I." means "confidence interval."

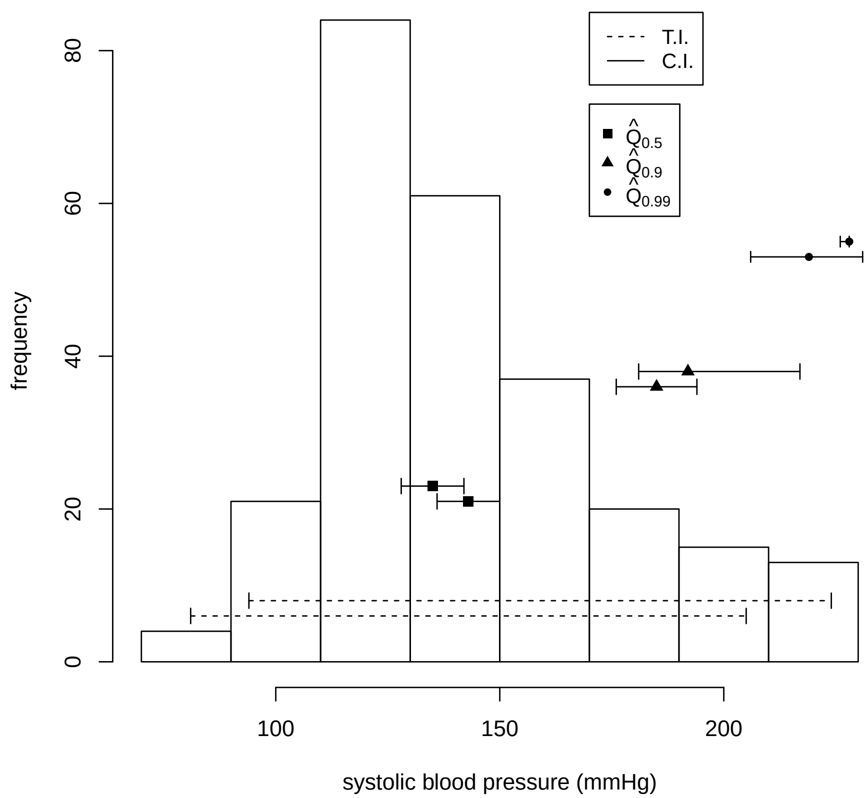


Figure 1: Histogram of the blood pressure data. Also marked on this graph are likelihood-based (bottom line segment) and nonparametric (top line segment) 95% confidence intervals for quantiles and $(0.90, 0.95)$ tolerance intervals. Here "T.I." means "tolerance interval" and "C.I." means "confidence interval." The measurements range between 77 to 228 mmHg.

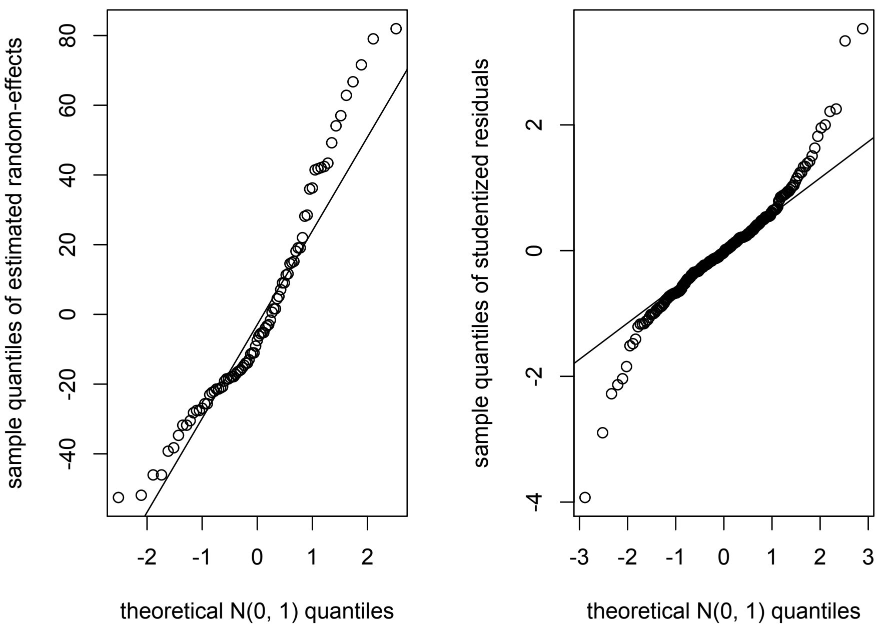


Figure 2: Normal quantile-quantile plots for estimated random effects and residuals resulting from fitting the model (27) to the blood pressure data. A line passing through the first and third quartiles is added in each plot.