# CONFIDENCE TUBES FOR MULTIPLE QUANTILE PLOTS VIA EMPIRICAL LIKELIHOOD

By John H. J. Einmahl[1] and Ian W. McKeague

*Eindhoven University of Technology and EURANDOM and Florida State University*

The nonparametric empirical likelihood approach is used to obtain simultaneous confidence tubes for multiple quantile plots based on $k$ independent (possibly right-censored) samples. These tubes are asymptotically distribution free, except when both $k \geq 3$ and censoring is present. Pointwise versions of the confidence tubes, however, are asymptotically distribution free in all cases. The various confidence tubes are valid under minimal conditions. The proposed methods are applied in three real data examples.

**1. Introduction.** The quantile-quantile (Q-Q) plot is a well known and attractive graphical method for comparing two distributions, especially when confidence limits are added. In this paper we develop Q-Q plot methods for the comparison of two or more distributions from randomly censored data. More specifically, we consider the problem of finding simultaneous confidence tubes for *multiple* quantile plots (for brevity, multi-Q plots) from $k$ independent samples of possibly right-censored survival times. The multi-Q plot is defined to be the $k$-dimensional curve $(Q_1(p), \ldots, Q_k(p))$ parameterized by $0 < p < 1$, where $Q_j$ is the quantile function of the $j$th distribution. It specializes to the ordinary Q-Q plot in the two-sample case.

The comparison of quantile functions is particularly useful for the analysis of survival data in biomedical settings. Frail and strong individuals (corresponding to low and high values of $p$) often respond to different treatments in different ways, so treatment effects can be hard to determine from comparison of mean or median survival times alone; see, for example, Doksum (1974). The approach developed here allows comparison of treatments simultaneously across all frailty levels.

Our approach is based on the nonparametric empirical likelihood method. This method was originally developed by Thomas and Grunkemeier (1975) and Owen (1988, 1990) as a way of improving upon Wald-type confidence regions. There now exists a substantial literature on empirical likelihood indicating that it is widely viewed as a desirable and natural approach to

statistical inference in a variety of settings. Moreover, there is considerable evidence that procedures based on the method outperform competing procedures. Empirical likelihood based confidence bands for individual quantile functions have recently been derived in Li, Hollander, McKeague and Yang (1996). Naik-Nimbalkar and Rajarshi (1997) employed the approach to test for equality of $k$ medians; their test naturally extends to a test for equality of $k$ quantiles.

We use the nonparametric empirical likelihood approach to derive asymptotic simultaneous confidence tubes for multi-Q plots based on $k$ independent random samples, including confidence bands for ordinary Q-Q plots ($k = 2$). The tubes are applicable to situations with or without random censoring. The limiting processes involved in the construction of the tubes are distribution free, except when $k \geq 3$ *and* censoring is present. In general, we are able to obtain asymptotically distribution-free pointwise confidence regions for the multi-Q plot. The various confidence tubes are valid under minimal conditions, although for convenience we shall assume continuity of the underlying distribution functions.

Q-Q plots are studied in detail using classical methods in Doksum (1974, 1977), Doksum and Sievers (1976) and Switzer (1976) for models without censoring; see Shorack and Wellner [(1986), pages 652–657] for a summary and discussion. For models with censoring, Wald-type simultaneous confidence bands for Q-Q plots are obtained in Aly (1986), but restrictive differentiability conditions on the underlying distribution functions are required. The $k$-sample problem without censoring is studied in Nair (1978, 1982), but essentially only pairwise comparisons are made there. A review of graphical methods in nonparametric statistics with extensive coverage of Q-Q plots can be found in Fisher (1983). Some refined approximation results for normalized Q-Q plots with statistical applications have been established in Beirlant and Deheuvels (1990) for the uncensored case and Deheuvels and Einmahl (1992) in the censored case.

The paper is organized as follows. The proposed confidence tubes and the main results are presented in Section 2. Our approach is illustrated in Section 3 using three real data examples. All the proofs are contained in Section 4.

**2. Main results.** We begin by specifying the setup precisely and introducing the basic notation. It is convenient first to recall the notation in the one-sample case. For the corresponding notation in the general $k$-sample case, we *use a further subscript j to refer to the jth sample.*

The random censorship model deals with $n$ i.i.d. pairs $(Z_i, \delta_i)$, $i = 1, \ldots, n$, obtained from two independent random samples $X_i$ and $Y_i$, $i = 1, \ldots, n$, in the following way: $Z_i = X_i \wedge Y_i$, $\delta_i = 1_{\{X_i \leq Y_i\}}$. The distribution functions of $X_i$ and $Y_i$ are denoted $F$ and $G$, respectively, and $F$ is assumed to be continuous. We will work with nonnegative $X_i$ and $Y_i$, but this restriction is in fact not needed anywhere; see the discussion at the end of this section. The (right-continuous) quantile function corresponding to $F$ is denoted by $Q$. We

write

$$L(\tilde{F}) = \prod_{i=1}^{n} \left(\tilde{F}(Z_i) - \tilde{F}(Z_i -)\right)^{\delta_i} \left(1 - \tilde{F}(Z_i)\right)^{1-\delta_i}$$

for the likelihood, where $\tilde{F}$ belongs to $\Theta$, the space of all distribution functions on $[0, \infty)$. The ordered uncensored survival times, that is, the $X_i$ with corresponding $\delta_i = 1$, are written $0 \le T_1 \le \cdots \le T_N < \infty$, and $r_j = \sum_{i=1}^{n} 1_{\{Z_i \ge T_j\}}$ denotes the size of the risk set at $T_j -$. The empirical likelihood ratio for $\tilde{F}(t) = p$ (given $0 < p < 1$) is defined by

$$R(t) = \frac{\sup\{L(\tilde{F}): \tilde{F}(t) = p, \tilde{F} \in \Theta\}}{\sup\{L(\tilde{F}): \tilde{F} \in \Theta\}}.$$

Note that the sup in the denominator is attained by the Kaplan–Meier (or product-limit) estimator

$$F_n(t) = 1 - \prod_{i: T_i \le t} \left(1 - \frac{1}{r_i}\right).$$

It can be shown with the aid of Lagrange's method [see Thomas and Grunkemeier (1975) or Li (1995)] that

$$-2 \log R(t) = -2 \sum_{i: T_i \le t} \left\{(r_i - 1)\log\left(1 + \frac{\lambda}{r_i - 1}\right) - r_k i \log\left(1 + \frac{\lambda}{r_i}\right)\right\},$$

where the Lagrange multiplier $\lambda > D := \max_{i: T_i \le t}(1 - r_i)$ satisfies the equation

$$(2.1) \qquad \prod_{i: T_i \le t} \left(1 - \frac{1}{r_i + \lambda}\right) = 1 - p.$$

Now we turn to the multisample setup. The $k$ samples are assumed to be independent with sample sizes denoted $n_1, \ldots, n_k$; write $n = \sum_{j=1}^{k} n_j$. Set $\mathbb{F} = (F_1, \ldots, F_k)$ and define the multi-Q plot to be

$$\{(Q_1(p), \ldots, Q_k(p)): 0 < p < 1\}.$$

Observe that this is the classical Q-Q plot when $k = 2$. In the sequel we consider the following more convenient version of the multi-Q plot: the graph **Q** of the function

$$t_1 \mapsto \left(Q_2(F_1(t_1)), \ldots, Q_k(F_1(t_1))\right)$$

for $t_1 \ge 0$. Denote the joint likelihood by

$$L(\tilde{\mathbb{F}}) = \prod_{j=1}^{k} L_j\left(\tilde{F}_j\right)$$

and the empirical likelihood ratio at $t = (t_1, \ldots, t_k)$ by

$$R(t) = \frac{\sup\left\{L(\tilde{\mathbb{F}}): \tilde{F}_j(t_j) = \tilde{F}_1(t_1) \text{ for all } j = 2, \ldots, k, \tilde{\mathbb{F}} \in \Theta^k\right\}}{\sup\{L(\tilde{\mathbb{F}}): \tilde{\mathbb{F}} \in \Theta^k\}}.$$

Again we find, using Lagrange's method with the $k - 1$ constraints, $\tilde{F}_1(t_1) = \tilde{F}_j(t_j)$, $j = 2, \ldots, k$, that

$$
\begin{aligned}
(2.2) \quad & -2 \log R(t) \\
& = -2 \sum_{j=1}^{k} \sum_{i:\, T_{ji} \le t_j} \left\{ (r_{ji} - 1) \log\left( 1 + \frac{\lambda_j}{r_{ji} - 1} \right) - r_{ji} \log\left( 1 + \frac{\lambda_j}{r_{ji}} \right) \right\},
\end{aligned}
$$

where the $\lambda_j$, $j = 2, \ldots, k$, satisfy the $k - 1$ equations

$$
(2.3) \qquad \prod_{i:\, T_{1i} \le t_1} \left( 1 - \frac{1}{r_{1i} + \lambda_1} \right) = \prod_{i:\, T_{ji} \le t_j} \left( 1 - \frac{1}{r_{ji} + \lambda_j} \right);
$$

here we have set $\lambda_1 = -\sum_{j=2}^{k} \lambda_j$ (so $\sum_{j=1}^{k} \lambda_j = 0$) and the $\lambda_j$ should satisfy $\lambda_j > D_j$ for $j = 1, \ldots, k$.

Later we show that this system of equations indeed has a unique solution; see Lemma 4.1. In the one-sample case, it is immediately clear that the corresponding Lagrange multiplier equation (2.1) has a unique solution, but it is not obvious in the multisample case. Computation of the $\lambda_j$'s can be carried out using a special-purpose root-finding procedure which exploits the monotonicity of the r.h.s. of (2.3) as a function of $\lambda_j$ (see Section 3 and the proof of Lemma 4.1).

The various confidence sets we propose are easily obtained from the main theorem below and are presented in the three subsequent theorems. These confidence sets are all of the form $\{\mathbf{t}\colon R(\mathbf{t}) > c\}$, where $c$ is derived using asymptotic considerations.

Before stating our main theorem we introduce some more notation. We assume throughout that $n_j/n \to p_j > 0$ as $n \to \infty$ for $j = 1, \ldots, k$ (although with some care this condition can be relaxed to $n_j \to \infty$). Define

$$
\sigma_j^2(s) = \int_0^s \frac{dF_j(u)}{\left( 1 - F_j(u) \right)\left( 1 - F_j(u-) \right)\left( 1 - G_j(u-) \right)}.
$$

We will need the $k \times k$-matrix $D = D(\mathbf{t})$ with entries

$$
d_{ij} = \begin{cases} -\dfrac{\sigma_i(t_i)\,\sigma_j(t_j)}{\sqrt{p_i\,p_j}}\,\eta_{ij}, & \text{for } j \ne i, \\[3mm] \dfrac{\sigma_i^2(t_i)}{p_i} \displaystyle\sum_{l \ne i} \eta_{li}, & \text{for } j = i, \end{cases}
$$

where

$$
\eta_{ij} = \frac{\prod_{l \ne i,\, l \ne j} \sigma_l^2(t_l)/p_l}{\sum_{q=1}^{k} \prod_{l \ne q} \sigma_l^2(t_l)/p_l}
$$

(the empty product is defined to be 1). Also define $V = V(\mathbf{t})$ to be the random $k$-vector with $j$th entry $W_j(\sigma_j^2(t_j))/\sigma_j(t_j)$, where the $W_j$ are independent standard Wiener processes. Let $\tau_1$ be such that $F_1(\tau_1) > 0$ and let $\tau_2 \ge \tau_1$ be

such that $F_1(\tau_2) < 1$, $G_1(\tau_2) < 1$ and $G_j(Q_j(F_1(\tau_2))) < 1$ for $j = 2, \ldots, k$. We assume throughout that the $F_j$ are continuous.

THEOREM 2.1. *When R, D and V are evaluated at* $\mathbf{t} = (t_1, Q_2(F_1(t_1)), \ldots, Q_k(F_1(t_1)))$ *for* $\tau_1 \le t_1 \le \tau_2$, *we have*

$$(2.4) \qquad\qquad -2 \log R \to_{\mathscr{D}} \|DV\|^2$$

*on* $D[\tau_1, \tau_2]$, *with* $\| \cdot \|$ *the k-dimensional Euclidian norm.*

Write the restriction of $\mathbf{Q}$ to $t_1 \in [\tau_1, \tau_2]$ as $\mathbf{Q}[\tau_1, \tau_2]$. In the next theorem we consider the important case $k = 2$, in which the multi-Q plot reduces to the usual Q-Q plot. Define $c_\alpha[s_1, s_2]$ for $0 < \alpha < 1$ by

$$P\left( \sup_{s \in [s_1, s_2]} W_1^2(s)/s < c_\alpha[s_1, s_2] \right) = 1 - \alpha.$$

Set $\hat{c}_\alpha = c_\alpha[\hat{\sigma}^2(\tau_1), \hat{\sigma}^2(\tau_2)]$, where

$$(2.5) \qquad \hat{\sigma}^2(t_1) = n\left\{ \frac{\hat{\partial}_1^2(t_1)}{n_1} + \frac{\hat{\sigma}_2^2\left(Q_{2n_2}(F_{1n_1}(t_1))\right)}{n_2} \right\},$$

with

$$(2.6) \qquad \hat{\sigma}_j^2(s) = n_j \sum_{i: T_{ji} \le s} \frac{1}{r_{ji}(r_{ji} - 1)}$$

and with $Q_{2n_2}$ the (right-continuous) quantile function corresponding to $F_{2n_2}$. Now we define the confidence band for $\mathbf{Q}[\tau_1, \tau_2]$ to be

$$\mathscr{B} = \left\{ \mathbf{t} \in [\tau_1, \tau_2] \times [0, \infty) : -2 \log R(t) < \hat{c}_\alpha \right\}.$$

THEOREM 2.2. *In the censored case, for* $k = 2$ *and* $0 < \alpha < 1$,

$$\lim_{n \to \infty} P(\mathbf{Q}[\tau_1, \tau_2] \in \mathscr{B}) = 1 - \alpha.$$

REMARK 2.1. In the uncensored case (and $k = 2$) we have that

$$\sigma^2(t_1) := \frac{\sigma_1^2(t_1)}{p_1} + \frac{\sigma_2^2(Q_2(F_1(t_1)))}{p_2}$$

$$(2.7) \qquad\qquad = \left( \frac{1}{p_1} + \frac{1}{p_2} \right) \frac{F_1(t_1)}{1 - F_1(t_1)}$$

$$= \left( \frac{1}{p_1} + \frac{1}{p_2} \right) \sigma_1^2(t_1).$$

Therefore for this case we can replace the $\hat{\sigma}^2(t_1)$ defined in (2.5) by the simpler but almost equivalent estimator

$$\hat{\sigma}^2(t_1) = n\left(\frac{1}{n_1} + \frac{1}{n_2}\right)\frac{F_{1n_1}(t_1)}{1 - F_{1n_1}(t_1)}.$$

For use in $\hat{c}_\alpha$, we can replace $\hat{\sigma}^2(t_1)$ by

$$\frac{F_{1n_1}(t_1)}{1 - F_{1n_1}(t_1)}.$$

Observe that this last expression is not an estimator of $\sigma^2(t_1)$ but of $\sigma_1^2(t_1)$. This however makes no difference because of (2.7) and the fact that for $c > 0$

$$\sup_{s \in [s_1, s_2]} W_1^2(s)/s =_{\mathscr{D}} \sup_{s \in [cs_1, cs_2]} W_1^2(s)/s.$$

Of course, here the Kaplan–Meier estimator $F_{1n_1}$ is just the empirical distribution function of the first sample.

Next we return to general $k \geq 2$, but assume that there is no censoring. Note that in this case the assumptions on $\tau_2$ reduce to $F_1(\tau_2) < 1$. Define $C_\alpha[s_1, s_2]$ for $0 < \alpha < 1$ by

$$(2.8) \qquad P\left(\sup_{s \in [s_1, s_2]} \frac{1}{s} \sum_{j=1}^{k-1} W_j^2(s) < C_\alpha[s_1, s_2]\right) = 1 - \alpha.$$

Set $\hat{C}_\alpha = C_\alpha[\hat{\sigma}_1^2(\tau_1), \hat{\sigma}_1^2(\tau_2)]$, where

$$\hat{\sigma}_1^2(t_1) = \frac{F_{1n_1}(t_1)}{1 - F_{1n_1}(t_1)}.$$

Define the confidence tube for $\mathbf{Q}[\tau_1, \tau_2]$ by

$$\mathscr{T} = \left\{\mathbf{t} \in [\tau_1, \tau_2] \times [0, \infty)^{k-1}: -2\log R(t) < \hat{C}_\alpha\right\}.$$

THEOREM 2.3.   *In the absence of censoring, for all $k \geq 2$ and $0 < \alpha < 1$,*

$$\lim_{n \to \infty} P(\mathbf{Q}[\tau_1, \tau_2] \in \mathscr{T}) = 1 - \alpha.$$

Now we allow censoring and $k \geq 2$ but take $\tau_2 = \tau_1$. Set $\mathbf{Q}[\tau_1] = \mathbf{Q}[\tau_1, \tau_1]$. Define the confidence region for $\mathbf{Q}[\tau_1]$ by

$$\mathscr{R} = \left\{\mathbf{t} \in \{\tau_1\} \times [0, \infty)^{k-1}: -2\log R(t) < \chi_\alpha^2\right\},$$

where $\chi_\alpha^2$ is the upper $\alpha$-quantile of the chi-square distribution with $k - 1$ degrees of freedom. In the case $k = 2$ note that $\mathscr{R}$ amounts to a confidence interval for the $F_1(\tau_1)$-quantile of $F_2$.

THEOREM 2.4. *In the censored case, for all $k \geq 2$ and $0 < \alpha < 1$,*

$$\lim_{n \to \infty} P(\mathbf{Q}[\tau_1] \in \mathscr{R}) = 1 - \alpha.$$

The asymptotic null distribution in the test for equality of $k$ medians developed by Naik-Nimbalkar and Rajarshi (1997) can be essentially derived from the proof of Theorem 2.4 by taking $\tau_1$ as their estimator $\theta^*$ of the common median.

Finally we establish an interval property for the confidence tube $\mathscr{T}$ (which also applies to $\mathscr{B}$ and $\mathscr{R}$): one-dimensional cross-sections parallel to a given axis are intervals. This is useful for computing the various confidence sets because points belonging to them can then be found by a simple search strategy that sweeps along each axis.

THEOREM 2.5. *Suppose that* $\mathbf{t}^{(l)} = (t_1, \ldots, t_j^{(l)}, \ldots, t_k) \in \mathscr{T}$ *for* $l = 1, 2$, *where* $t_j^{(1)} < t_j^{(2)}$. *Then* $\mathbf{t}^* = (t_1, \ldots, t_j^*, \ldots, t_k) \in \mathscr{T}$ *for any* $t_j^* \in [t_j^{(1)}, t_j^{(2)}]$.

In the two-sample case ($k = 2$) we have a somewhat stronger result.

THEOREM 2.6. *Let* $(t_1^{(l)}, t_2^{(l)})$, $l = 1, 2$, *belong to the confidence band* $\mathscr{B}$ *and suppose* $t_1^{(1)} \geq t_1^{(2)}$ *and* $t_2^{(2)} \geq t_2^{(1)}$. *Then* $(t_1^*, t_2^*)$ *also belongs to* $\mathscr{B}$ *whenever* $t_1^{(2)} \leq t_1^* \leq t_1^{(1)}$ *and* $t_1^{(1)} \leq t_2^* \leq t_2^{(2)}$.

This theorem (as well as Theorem 2.5) implies, by taking $t_1^{(1)} = t_1^{(2)}$ or $t_2^{(1)} = t_2^{(2)}$, that the intersection of the band $\mathscr{B}$ with a vertical or horizontal line is an interval. In addition, it shows that the bands are nondecreasing in the sense that their lower or upper boundaries are nondecreasing.

*Discussion.* We wish to emphasize that our approach, including the definition of the multi-Q plot, is new even in the *uncensored* case. We also remind that nonnegativity of the observations is not needed anywhere in the proofs. This is especially useful in the uncensored case, where often the $k$ samples do not represent life or failure times or when a transformation is applied to the data (see the third example in Section 3).

Another desirable feature of our approach is that the confidence bands and tubes are essentially invariant under permutations of the order of the $k$ samples involved. (Only at the two "ends" of the tube does the first sample play a somewhat special role.)

We did not formulate a version of our confidence tubes in the censored case for $k \geq 3$ since then $\|DV\|^2$ in (2.4) is not distribution free, even when only one of the $k$ samples is subject to censoring. Our approach can, however, be generalized to this situation by estimating all the unknowns appearing in $D$ and $V$ and then using simulation. This means that we replace $D$ by $C$ (given in the proof of Theorem 2.1), $t_j$ by $Q_{jn_j}(F_{1n_1}(t_1))$ for $j = 2, \ldots, k$, and $\sigma_j^2$ by $\hat{\sigma}_j^2$. The process to be simulated has the form $\|CV\|^2$, where $\hat{V}$ is the estimated version of $V$. Hence, approximate $1 - \alpha$ confidence tubes can be

constructed for the censored case as well, but we do not pursue this in further detail here.

The one-sample Q-Q plot, $t \mapsto Q(F_0(t))$ with $F_0$ known, is essentially treated in Li, Hollander, McKeague and Yang (1996), since their confidence bands for $Q(p)$ can be transformed to bands for $Q(F_0(t))$ by the time change $p = F_0(t)$. The present paper can be seen as a generalization of their approach to the $k$-sample case.

For uncensored data, in the two-sample Q-Q plot case, our confidence bands perform well in the tails due to the weighting which naturally arises when using the empirical likelihood method. Our bands share this property with the weighted bands (W bands) introduced in Doksum and Sievers (1976), which are based on the standardized two-sample empirical process. The bands in Switzer (1976) [and Aly (1986) for the censored case] are much wider in the tails, since they are based on the *unweighted* empirical process. All these procedures as well as our procedures are essentially based on the inversion of a distance between empirical distribution functions (or Kaplan–Meier estimators). In fact, the W bands are asymptotically equivalent to our bands in the uncensored case.

## 3. Applications to real data.

In this section we illustrate our approach in three real data examples.

First we consider a biomedical example for the two-sample case with censored survival data. The data come from a Mayo Clinic trial involving a treatment for primary biliary cirrhosis of the liver; see Fleming and Harrington (1991) for discussion. A total of $n = 312$ patients participated in the randomized clinical trial, 158 receiving the treatment (D-penicillamine) and 154 receiving a placebo. Censoring is heavy (187 of the 312 observations are censored). Figure 1 displays the 90% confidence band (and pointwise confidence intervals) for the Q-Q plot of treatment versus placebo for survival time in days. The standard empirical Q-Q plot based on quantiles of the Kaplan–Meier estimator is also displayed. Note that although the diagonal departs from the pointwise confidence region at some points, it remains within the simultaneous band, so there is no overall evidence of a difference between treatment and placebo.

The second example also illustrates the two-sample case. Hollander, McKeague and Yang (1997) analyzed data on 432 manuscripts submitted to the Theory and Methods Section of *JASA* during 1994. Each observation consists of the number of days between a manuscript's submission and its first review or the end of the year, along with a censoring indicator (1 if a paper received its first review by the end of the year; 0 otherwise). Similar data (on 444 manuscripts) are available for 1995. The censoring is light (330 of the 876 observations are censored) compared with the previous example. It is of interest to look for differences in the pattern of review times for the two years. Figure 2 displays the 95% confidence band (and pointwise confidence intervals) for the Q-Q plot. The lower endpoints of the pointwise confidence intervals touch the diagonal between 10 and 25 days, which might suggest
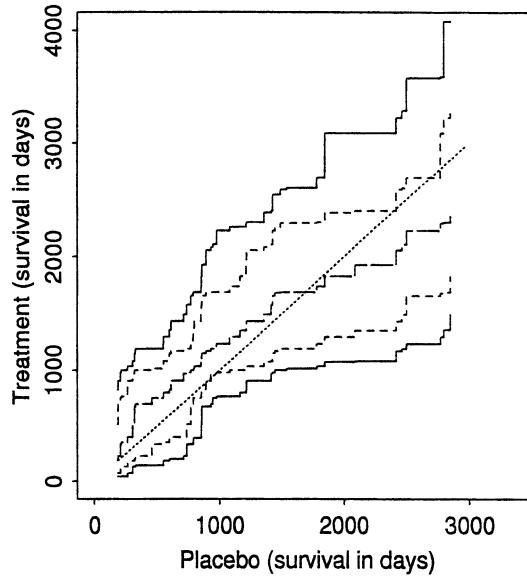
FIG. 1.  90% *confidence band* (*solid line*) *for the treatment versus placebo Q-Q plot in the Mayo Clinic trial, for* $186 \leq t_1 \leq 2976$ *days; pointwise confidence intervals* (*short dashed line*), *empirical Q-Q plot* (*long dashed line*).

that "rapid" reviews were faster in 1994 than in 1995. However, the diagonal is completely contained within the simultaneous band, so there is no overall evidence of a difference between the patterns of review times.

The third example concerns times to breakdown (in minutes) of an insulating fluid under three elevated voltage stresses, from data reported in Nair (1982), Table 1. It is important to determine whether the distribution of time to breakdown changes with voltage. There are 60 uncensored observations at each voltage level (34, 35 and 36 Kv). As in Nair (1982) we use the 34 Kv measurements as a reference sample and put the breakdown times on a log-scale. Figure 3 shows cross-sections of the 95% confidence regions for the multi-Q plot at three values of the reference sample: $t_1 = 0.41$, $1.06$ and $1.65$. The confidence tube gives simultaneous coverage over the interval $0.41 \leq t_1 \leq 1.65$. The diagonal $(t_1, t_1, t_1)$ runs above the pointwise confidence region at $t_1 = 1.65$ (top right plot) suggesting that increased voltage can reduce breakdown time in the upper tail of the distribution. However, the diagonal falls completely inside the simultaneous tube (left column) so there is a lack of significant evidence for breakdown time changing with voltage.

In these examples, we computed the Lagrange multipliers in the system of equations (2.3) using the van Wijngaarden–Dekker–Brent root finding algorithm [Press, Teukolsky, Vetterling and Flannery (1992), page 359]. The proof of Lemma 4.1 provides a constructive method to obtain the solution by repeated use of their algorithm. The thresholds $\hat{c}_\alpha$ and $\hat{C}_\alpha$ used in the
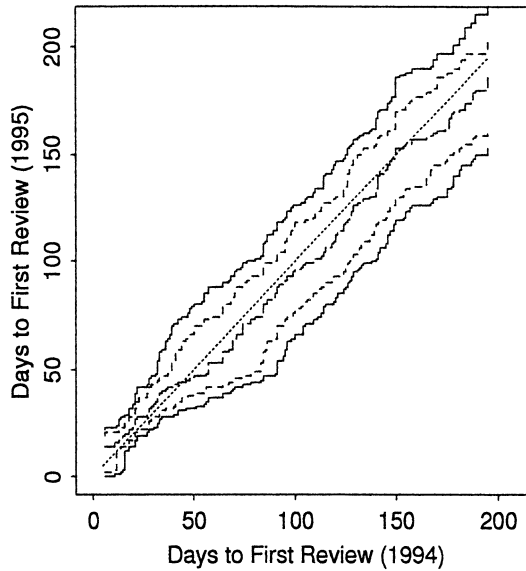
FIG. 2. *95% confidence band (solid line) for the Q-Q plot based on the JASA time-to-first-review data, for $5 \le t_1 \le 195$ days; pointwise confidence intervals (short dashed line), empirical Q-Q plot (long dashed line).*

confidence bands and tubes were computed by simulation of the Wiener processes on a fine grid.

**4. Proofs.** Here we present proofs of the theorems in Section 2. Some lemmas used in these proofs are given at the end of this section.

PROOF OF THEOREM 2.1. First we note that, by Lemma 4.1, below the system of equations in (2.3) with $\lambda_1$ replaced by $-\sum_{j=2}^{k} \lambda_j$ has a unique solution for all $k \ge 2$.

Define $g_j: (D_j, \infty) \to \mathbb{R}$ by

$$g_j(\lambda) = \sum_{i:\, T_{ji} \le t_j} \log\left(1 - \frac{1}{r_{ji} + \lambda}\right)$$

for $j = 1, \ldots, k$. Denote $a_j = g_j(0) = \log \hat{S}_j(t_j)$, where $\hat{S}_j$ is the Kaplan–Meier estimator of $S_j = 1 - F_j$ and $b_j = g_j'(0) = \hat{\sigma}_j^2(t_j)/n_j$ with $\hat{\sigma}_j^2$ as in (2.6). Here $t_j = Q_j(F_1(t_1))$ for $j = 2, \ldots, k$. Taylor series expansions of $g_1$ and $g_j$, in conjunction with Lemma 4.2 and the argument of Li [(1995), proof of (2.15), page 102] yield

$$(4.1) \qquad 0 = g_1(\lambda_1) - g_j(\lambda_j) = a_1 - a_j + \lambda_1 b_1 - \lambda_j b_j + O_P(n^{-1}),$$
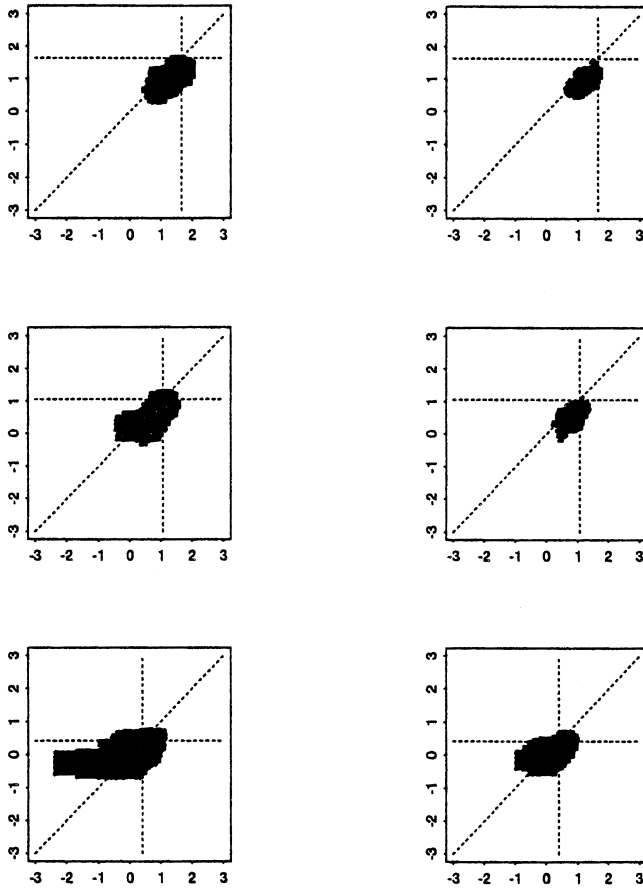
FIG. 3.  *Time to insulating fluid breakdown* (*in log-scale*), 36 *Kv sample versus* 35 *Kv sample*; *cross-sections of the* 95% *simultaneous confidence tube* (*left column*) *and pointwise confidence regions* (*right column*) *at* $t_1 = 0.41$, 1.06 *and* 1.65 *in the* 34 *Kv reference sample* (*bottom row to top row*, *respectively*).

uniformly in $t_1 \in [\tau_1, \tau_2]$. Ignore the remainder term for the moment and consider the system of equations

$$(4.2) \qquad \begin{aligned} \tilde{\lambda}_j b_j - \tilde{\lambda}_1 b_1 &= a_1 - a_j \quad \text{for } j = 2, \ldots, k, \\ \tilde{\lambda}_1 + \cdots + \tilde{\lambda}_k &= 0 \end{aligned}$$

with unknowns $\tilde{\lambda}_1, \ldots, \tilde{\lambda}_k$. By Lemma 4.3 this system has as unique solution $\tilde{\lambda}_j = \sum_{i \neq j} (a_i - a_j) \gamma_{ij}$ with the $\gamma_{ij}$ as defined in the lemma. We now use this result to obtain an approximation for the $\lambda_j$.

The remainder term in (4.1) consists of the remainders in the Taylor series expansions of $g_1$ and $g_j$, and both are of order $O_p(n^{-1})$. Attach these remainder terms to $a_1$ and $a_j$, respectively, and apply Lemma 4.3. Note that

$\hat{\sigma}_j^2$ is a uniformly consistent estimator of $\sigma_j^2$, so $b_j = O_P(n^{-1})$ and $\gamma_{ij} = O_P(n)$, and it follows that

$$\lambda_j = \tilde{\lambda}_j + O_P(n^{-1})O_P(n) = \tilde{\lambda}_j + O_P(1),$$

uniformly for $t_1 \in [\tau_1, \tau_2]$. We also have that

(4.3) $$\lambda_j^2 b_j = \tilde{\lambda}_j^2 b_j + O_P(n^{-1/2}).$$

Applying the Taylor series argument of Li [(1995), page 102] to (2.2) and using (4.3) then gives

$$-2 \log R(t) = \sum_{j=1}^{k} \tilde{\lambda}_j^2 b_j + O_P(n^{-1/2}).$$

Write the leading term above in the form

$$\sum_{j=1}^{k} \tilde{\lambda}_j^2 b_j = \|Cw\|^2,$$

where $C$ is the $k \times k$-matrix with entries

$$c_{ij} = \begin{cases} -\sqrt{b_i b_j}\, \gamma_{ij} & \text{for } j \neq i, \\ b_i \sum_{l \neq i} \gamma_{li}, & \text{for } j = i. \end{cases}$$

and $w$ is the $k$-vector with entries

$$w_j = \frac{a_j - \log S_j(t_j)}{\sqrt{b_j}} = \frac{a_j - \log S_1(t_1)}{\sqrt{b_j}}.$$

The proof is completed by noting that

$$\left(w_j(t_j)\right)_{j=1,\ldots,k} \to_{\mathscr{D}} \left(\frac{W_j\left(\sigma_j^2(t_j)\right)}{\sigma_j(t_j)}\right)_{j=1,\ldots,k} = V(t_1),$$

where $W_1, \ldots, W_k$ are independent standard Wiener processes, and $c_{ij} \to_P d_{ij}$.
$\square$

PROOF OF THEOREM 2.2. Let us first simplify $\|DV\|^2$ for this case. Note that for $k = 2$ we have

$$D = \frac{1}{\sigma^2(t_1)} \begin{pmatrix} \dfrac{\sigma_1^2(t_1)}{p_1} & \dfrac{-\sigma_1(t_1)\sigma_2(t_2)}{\sqrt{p_1 p_2}} \\ \dfrac{-\sigma_1(t_1)\sigma_2(t_2)}{\sqrt{p_1 p_2}} & \dfrac{\sigma_2^2(t_2)}{p_2} \end{pmatrix},$$

with $\sigma^2$ as in Remark 2.1. So

$$
D = \frac{1}{\sigma^2(t_1)}
\begin{pmatrix}
\dfrac{\sigma_1(t_1)}{\sqrt{p_1}} \\[2ex]
\dfrac{-\sigma_2(t_2)}{\sqrt{p_2}}
\end{pmatrix}^{\otimes 2},
$$

where $a^{\otimes 2} = aa'$, and hence

$$
DV = \frac{1}{\sigma^2(t_1)}
\begin{pmatrix}
\dfrac{\sigma_1(t_1)}{\sqrt{p_1}} \\[2ex]
\dfrac{-\sigma_2(t_2)}{\sqrt{p_2}}
\end{pmatrix}
\left( \frac{W_1\big(\sigma_1^2(t_1)\big)}{\sqrt{p_1}} - \frac{W_2\big(\sigma_2^2(t_2)\big)}{\sqrt{p_2}} \right)
$$

$$
=_{\mathscr{D}} \frac{1}{\sigma^2(t_1)}
\begin{pmatrix}
\dfrac{\sigma_1(t_1)}{\sqrt{p_1}} \\[2ex]
\dfrac{-\sigma_2(t_2)}{\sqrt{p_2}}
\end{pmatrix}
W_1\big(\sigma^2(t_1)\big),
$$

so that

$$
\| DV \|^2 =_{\mathscr{D}} \frac{W_1^2\big(\sigma^2(t_1)\big)}{\sigma^2(t_1)}.
$$

It is well known that $\hat{\sigma}_j^2(s) \to_P \sigma_j^2(s)$, $j = 1, 2$, and hence with some care it can be shown that $\hat{\sigma}^2(\tau_l) \to_P \sigma^2(\tau_l)$, $l = 1, 2$. Setting $c_\alpha = c_\alpha[\sigma^2(\tau_1), \sigma^2(\tau_2)]$, this yields $\hat{c}_\alpha \to_P c_\alpha$.

Combining the above we obtain

$$
P\big(\mathbf{Q}[\tau_1, \tau_2] \subset \mathscr{B}\big) = P\big(-2 \log R\big(t_1, Q_2\big(F_1(t_1)\big) < \hat{c}_\alpha \text{ for all } t_1 \in [\tau_1, \tau_2]\big)
$$

$$
\to P\left( \sup_{t_1 \in [\tau_1, \tau_2]} \frac{W_1^2\big(\sigma^2(t_1)\big)}{\sigma^2(t_1)} < c_\alpha \right)
$$

$$
= P\left( \sup_{s \in [\sigma^2(\tau_1), \sigma^2(\tau_2)]} \frac{W_1^2(s)}{s} < c_\alpha \right) = 1 - \alpha,
$$

where we used, for the convergence statement, that the random variable in the last expression has a continuous distribution. $\square$

Before continuing with the proofs of the theorems let us do some calculations on $\| DV \|^2$ of Theorem 2.1 in general. Note that $D$ is symmetric and by

Lemma 4.4 it is idempotent of rank $k - 1$. Thus we may diagonalize $D = D(t_1)$ as follows:

$$(4.4) \qquad D(t_1) = P(t_1)' \begin{pmatrix} I_{k-1} & 0 \\ 0 & 0 \end{pmatrix} P(t_1),$$

where $P(t_1)$ is orthogonal and $I_{k-1}$ is the identity matrix of order $k - 1$. Put $Z(t_1) = P(t_1)V(t_1)$. Then

$$(4.5) \qquad \begin{aligned} \|D(t_1)V(t_1)\|^2 &= V(t_1)'D(t_1)'D(t_1)V(t_1) = V(t_1)'D(t_1)V(t_1) \\ &= Z(t_1)' \begin{pmatrix} I_{k-1} & 0 \\ 0 & 0 \end{pmatrix} Z(t_1), \end{aligned}$$

where the second equality follows since $D$ is symmetric and idempotent. The covariance structure of the process $Z(t_1)$ is given, for two values of $t_1$, say $s \le t$, by

$$(4.6) \qquad \begin{aligned} &E(Z(s)Z(t)') \\ &= P(s)\mathrm{diag}\left( \frac{\sigma_1(s)}{\sigma_1(t)}, \frac{\sigma_2(Q_2(F_1(s)))}{\sigma_2(Q_2(F_1(t)))}, \ldots, \frac{\sigma_k(Q_k(F_1(s)))}{\sigma_k(Q_k(F_1(t)))} \right) P(t)'. \end{aligned}$$

PROOF OF THEOREM 2.3.   First observe that

$$\sigma_j^2(Q_j(F_1(t_1))) = \frac{F_j(Q_j(F_1(t_1)))}{1 - F_j(Q_j(F_1(t_1)))} = \frac{F_1(t_1)}{1 - F_1(t_1)} = \sigma_1^2(t_1),$$

for $j = 2, \ldots, k$. This implies $D(t_1)$, and hence $P(t_1)$, does not depend on $t_1$. Thus the r.h.s. of (4.6) reduces to

$$\frac{\sigma_1(s)}{\sigma_1(t)} I_k.$$

It follows that the process $Z(t_1)$ has the same distribution as the process

$$\left( \frac{W_1(\sigma_1^2(t_1))}{\sigma_1(t_1))}, \ldots, \frac{W_k(\sigma_1^2(t_1))}{\sigma_1(t_1)} \right)',$$

where the $W_j$'s are independent standard Wiener processes, and hence by (4.5),

$$\|DV\|^2 =_{\mathscr{D}} \sum_{j=1}^{k-1} \frac{W_j^2(\sigma_1^2(t_1))}{\sigma_1^2(t_1)}.$$

Now the proof of this theorem can be completed along the same lines as that of Theorem 2.2. In this case, use continuity of the random variable

$$\sup_{s \in [\sigma_1^2(\tau_1), \sigma_1^2(\tau_2)]} \sum_{j=1}^{k-1} \frac{W_j^2(s)}{s},$$

which follows from a property of Gaussian measures on Banach spaces, namely that the measure of a closed ball is a continuous function of its radius; apply, for example, Paulauskas and Račkauskas [(1989), Chapter 4, Theorem 1.2] to the Gaussian measure induced by the process $s^{-1/2}(W_1(s), \ldots, W_{k-1}(s))$ on the Banach space of $\mathbb{R}^{k-1}$-valued continuous functions on $[\sigma_1^2(\tau_1), \sigma_1^2(\tau_2)]$ endowed with the supremum norm. $\square$

PROOF OF THEOREM 2.4.  This theorem can be proved along the lines of the previous two. We only note that now the r.h.s. of (4.6), with $s = t = \tau_1$, reduces to the identity matrix $I_k$. Thus $Z(\tau_1)$ is a $k$-vector of independent standard normal random variables. Hence from (4.5) we find that $\|DV\|^2$, evaluated at $\tau_1$, has a $\chi_{k-1}^2$ distribution. $\square$

PROOF OF THEOREM 2.5.  In order not to overdo the notation we restrict ourselves to proving this theorem for $k = 3$; for $k \neq 3$ the proof is essentially the same. W.l.o.g. we take $j = 3$. Because the denominator of the likelihood ratio does not depend on $\mathbf{t} = (t_1, \ldots, t_k)$, we only consider the expression

$$-2 \log \prod_{j=1}^{3} \prod_{i=1}^{N_j} h_{ji}(1 - h_{ji})^{r_{ji}-1},$$

with the $h_{ji} \in (0, 1)$ defined by $h_{ji} = \tilde{F}_j(T_{ji})/(1 - \tilde{F}_j(T_{j,i-1}))$. Setting $z_{ji} = \log(1 - h_{ji})$, this becomes

$$-2 \log \prod_{j=1}^{3} \prod_{i=1}^{N_j} (1 - \exp(z_{ji})) \exp(z_{ji}(r_{ji} - 1))$$

$$= -2 \sum_{j=1}^{3} \sum_{i=1}^{N_j} \left\{ z_{ji}(r_{ji} - 1) + \log(1 - \exp(z_{ji})) \right\} =: g(\mathbf{z}),$$

with $\mathbf{z} = (z_{11}, \ldots, z_{1N_1}, z_{21}, \ldots, z_{2N_2}, z_{31}, \ldots, z_{3N_3})$. Observe that $g$ is a convex function.

Now $g(\mathbf{z})$, $\mathbf{z} \in (-\infty, 0)^{N_1+N_2+N_3}$, has to be minimized under the constraints

$$(4.7) \qquad \sum_{i:\, T_{1i} \leq t_1} z_{1i} = \sum_{i:\, T_{2i} \leq t_2} z_{2i} \quad \text{and} \quad \sum_{i:\, T_{1i} \leq t_1} z_{1i} = \sum_{i:\, T_{3i} \leq t_3} z_{3i}.$$

Solutions of (4.7) for $\mathbf{t} = \mathbf{t}^{(l)}$ that minimize $g(\mathbf{z})$, are denoted with $\mathbf{z}^{(l)}$, $l = 1, 2$, respectively.

For $t_3^* \in [t_3^{(1)}, t_3^{(2)}]$, define the function

$$f(x) = \sum_{i:\, T_{1i} \leq t_1} \left( x z_{1i}^{(1)} + (1 - x) z_{1i}^{(2)} \right) - \sum_{i:\, T_{3i} \leq t_3^*} \left( x z_{3i}^{(1)} + (1 - x) z_{3i}^{(2)} \right)$$

for $0 \leq x \leq 1$. Since $t_3^* \leq t_3^{(2)}$, we easily see that $f(0) \leq 0$. Similarly, using $t_3^* \geq t_3^{(1)}$, we obtain $f(1) \geq 0$. Thus there exists an $x' \in [0, 1]$ such that $f(x^*) = 0$.

Define

$$\mathbf{z}^* = \left( x^* z_{11}^{(1)} + (1 - x^*) z_{11}^{(2)}, \ldots, x^* z_{3N_3}^{(1)} + (1 - x^*) z_{3N_3}^{(2)} \right).$$

Then trivially the two equations in (4.7) are satisfied for $z = \mathbf{z}^*$ and $t = \mathbf{t}^*$. Also because $g$ is convex,

$$g(\mathbf{z}^*) \leq x^* g(\mathbf{z}^{(1)}) + (1 - x^*) g(\mathbf{z}^{(2)}).$$

This implies, since $-2 \log R(\mathbf{t}^{(l)}) < \hat{C}_\alpha$, $l = 1, 2$, that $-2 \log R(\mathbf{t}^*) < \hat{C}_\alpha$, that is, $\mathbf{t}^* \in \mathscr{T}$. □

The proof of Theorem 2.6 is similar to, but easier than, the previous proof. Moreover it is a straightforward extension of the proof of Theorem 1 in Li, Hollander, McKeague and Yang (1996). Therefore we will omit the proof here.

We conclude by proving the four lemmas that we used earlier.

LEMMA 4.1. *The system of equations* (2.3), *with unknowns* $\lambda_2, \ldots, \lambda_k$, *has a unique solution for all* $k \geq 2$ *provided* $D_j < 0$ *for* $j = 1, \ldots, k$.

PROOF. Define $f_j \colon (D_j, \infty) \to (0, 1)$ by

$$(4.8) \qquad f_j(\lambda) = \prod_{i : T_{ji} \leq t_j} \left(1 - \frac{1}{r_{ji} + \lambda}\right)$$

for $j = 1, \ldots, k$. We need to show that the system of equations

$$(4.9) \qquad f_1\left(-\sum_{j=2}^{k} \lambda_j\right) = f_j(\lambda_j), \qquad j = 2, \ldots, k$$

has a unique solution. Note that $f_j$ is continuous, strictly increasing, and vanishes as $\lambda_j \downarrow D_j$. It then follows that there is a unique solution to (4.9) when $k = 2$, because the decreasing function $f_1(-\lambda_2)$ must cross the increasing function $f_2(\lambda_2)$ at exactly one value of $\lambda_2 \in (D_2, -D_1)$.

Now consider $k \geq 3$. For each fixed $\lambda_2 > D_2$ and $j = 3, \ldots, k$, there exists a unique $\lambda_j = \lambda_j(\lambda_2)$ such that $f_2(\lambda_2) = f_j(\lambda_j)$. Each of these $\lambda_j$'s is strictly increasing as a function of $\lambda_2$ because $f_2$ and $f_j$ are strictly increasing. Now consider the equation

$$(4.10) \qquad f_1\left(-\lambda_2 - \sum_{j=3}^{k} \lambda_j(\lambda_2)\right) = f_2(\lambda_2).$$

The l.h.s. of (4.10) is defined whenever $D_2 < \lambda_2 < D_2^*$, where $D_2^*$ is the unique solution to

$$-D_2^* - \sum_{j=3}^{k} \lambda_j(D_2^*) = D_1.$$

Note that $D_2 < D_2^*$ because

$$-D_2 - \sum_{j=3}^{k} \lambda_j(D_2) = -\sum_{j=2}^{k} D_j > 0 > D_1.$$

Moreover, as a function of $\lambda_2 \in (D_2, D_2^*)$, the l.h.s. of (4.10) is strictly decreasing and vanishes as $\lambda_2 \uparrow D_2^*$; the r.h.s. is strictly increasing and

vanishes as $\lambda_2 \downarrow D_2$. Thus (4.10) holds for some unique $\lambda_2 = \bar{\lambda}_2 \in (D_2, D_2^*)$. Now set $\bar{\lambda}_j = \lambda_j(\bar{\lambda}_2)$ for $j = 3, \ldots, k$. It is then clear that $(\bar{\lambda}_2, \ldots, \bar{\lambda}_k)$ is the unique solution to (4.9). $\square$

LEMMA 4.2.  *Suppose* $n_j/n \to p_j > 0$ *for* $j = 1, \ldots, k$. *Let* $t_j = Q_j(F_1(t_1))$ *for* $j = 2, \ldots, k$ *and* $\mathbf{t} = (t_1, \ldots, t_k)$. *Then*

$$\lambda_j = \lambda_j(t_1) = O_P(n^{1/2}) \text{ uniformly over } [\tau_1, \tau_2].$$

PROOF.   Write the value of each side of (2.3) as $1 - p$ when $\mathbf{t}$ has the above form. By Li [(1995), page 101], if $\lambda_j < 0$ then

$$-\log(1 - p) \geq \hat{A}_j(t_j)\left( \frac{n_j}{n_j + \lambda_j} \right),$$

where $\hat{A}_j$ is the Nelson–Aalen estimator of $A_j$, the cumulative hazard function corresponding to $F_j$, and if $\lambda_j \geq 0$ then the above inequality reverses. Thus for any pair $\lambda_j, \lambda_l$ with $\lambda_j < 0$, $\lambda_l \geq 0$ (such pairs always exist, if not all the $\lambda_j$'s are 0, since $\sum_{j=1}^{k} \lambda_j = 0$) we have

$$\hat{A}_j(t_j)\left( \frac{n_j}{n_j + \lambda_j} \right) \leq \hat{A}_l(t_l)\left( \frac{n_l}{n_l + \lambda_l} \right),$$

and hence

$$\lambda_l n_j \hat{A}_j(t_j) - \lambda_j n_l \hat{A}_l(t_l) \leq \left( \hat{A}_l(t_l) - \hat{A}_j(t_j) \right) n_l n_j.$$

Note that $A_j(t_j) = A_1(t_1)$ and $A_1(t_1)$ is bounded away from 0 if $t_1 \geq \tau_1$. Thus by the uniform convergence of the Nelson–Aalen estimators $\hat{A}_j$, we have that for any $\varepsilon > 0$ and $n$ sufficiently large, $\hat{A}_j(t_j) \geq \frac{1}{2} A_1(t_1)$ for all $t_1 \in [\tau_1, \tau_2]$ with probability at least $1 - \varepsilon$, similarly for $\hat{A}_l$. It then follows that

$$0 \leq \tfrac{1}{2}\big(\lambda_l n_j - \lambda_j n_l\big) A_1(t_1) \leq \left( \hat{A}_l(t_l) - \hat{A}_j(t_j) \right) n_l n_j,$$

with probability $1 - \varepsilon$, for $n$ sufficiently large. Finally, using the fact that $\hat{A}_j(t_j) = A_1(t_1) + O_P(n^{-1/2})$ uniformly over $[\tau_1, \tau_2]$, we find that $\lambda_j = O_P(n^{1/2})$ for all $j = 1, \ldots, k$, uniformly for $t_1 \in [\tau_1, \tau_2]$. $\square$

LEMMA 4.3.   *The system of equations* (4.2) *has solution*

$$\tilde{\lambda}_j = \sum_{i \neq j} (a_i - a_j)\gamma_{ij},$$

*where*

$$\gamma_{ij} = \gamma_0 \prod_{l \neq i,\, l \neq j} b_l \quad and \quad \gamma_0 = \left( \sum_{i=1}^{k} \prod_{l \neq i} b_l \right)^{-1}.$$

*The solution is unique when all the $b_l$'s are positive.*

PROOF. The coefficient of $a_1$ in $\sum_{j=1}^{k} \tilde{\lambda}_j$ is

$$\sum_{i \neq 1} \gamma_{i1} - \sum_{j \neq 1} \gamma_{1j} = 0,$$

similarly for the coefficients of $a_2, \ldots, a_k$. Thus $\sum_{j=1}^{k} \tilde{\lambda}_j = 0$. The coefficient of $a_1$ in $\tilde{\lambda}_2 b_2$, is

$$b_2 \gamma_{12} = \gamma_0 \prod_{l \neq 1} b_l$$

and the coefficient of $a_1$ in $\tilde{\lambda}_1 b_1$ is

$$-b_1 \sum_{i \neq 1} \gamma_{i1} = -\gamma_0 \sum_{i \neq 1} \prod_{l \neq i} b_l$$

so the coefficient of $a_1$ in $\tilde{\lambda}_2 b_2 - \tilde{\lambda}_1 b_1$ is

$$\gamma_0 \prod_{l \neq 1} b_l + \gamma_0 \sum_{i \neq 1} \prod_{l \neq i} b_l = 1.$$

The same argument shows that the coefficient of $a_2$ in $\tilde{\lambda}_2 b_2 - \tilde{\lambda}_1 b_1$ is $-1$. The coefficient of $a_q$, with $q \geq 3$, in $\tilde{\lambda}_2 b_2 - \tilde{\lambda}_1 b_1$ is

$$b_2 \gamma_{q2} - b_1 \gamma_{q1} = \gamma_0 \left( b_2 \prod_{l \neq 2,\, l \neq q} b_l - b_1 \prod_{l \neq 1,\, l \neq q} b_l \right) = 0.$$

This shows that $\tilde{\lambda}_2 b_2 - \tilde{\lambda}_1 b_1 = a_1 - a_2$ and the same argument shows that all the other equations in (4.2) are satisfied. $\square$

LEMMA 4.4. *The $k \times k$-matrix $D = D(\mathbf{t})$ is idempotent, that is, $D^2 = D$ and of rank $k - 1$.*

PROOF. Setting $v_j = \sigma_j^2(t_j)/p_j$, we have

$$d_{ii} = \frac{\sum_{j \neq i} \prod_{l \neq j} v_l}{\sum_{q=1}^{k} \prod_{l \neq q} v_l},$$

$$d_{ij} = \frac{-\sqrt{v_i v_j} \prod_{l \neq i,\, l \neq j} v_l}{\sum_{q=1}^{k} \prod_{l \neq q} v_l}, \qquad i \neq j.$$

Because of the various symmetries it suffices to show that

(4.11) $$d_{11} = \sum_{i=1}^{k} d_{1i}^2 \quad \text{and} \quad d_{12} = \sum_{i=1}^{k} d_{1i}, d_{2i},$$

for the idempotency of $D$. For the first equality we need to show that

$$\left( \sum_{j=2}^{k} \prod_{l \neq j} v_l \right) \left( \sum_{i=1}^{k} \prod_{l \neq i} v_l \right) = \left( \sum_{j=2}^{k} \prod_{l \neq j} v_l \right)^2 + \sum_{j=2}^{k} v_1 v_j \left( \prod_{l \neq 1,\, l \neq j} v_l \right)^2.$$

Writing $C = \sum_{j=2}^{k} \prod_{l \neq j} v_l$, this reduces to

$$C\left(C + \prod_{l \neq 1} v_l\right) = C^2 + v_1 \sum_{j=2}^{k} v_j \left(\prod_{l \neq 1, l \neq j} v_l\right)^2,$$

or, subtracting $C^2$ on both sides,

$$\left(\sum_{j=2}^{k} \prod_{l \neq j} v_l\right) \prod_{l \neq 1} v_l = v_1 \sum_{j=2}^{k} v_j \left(\prod_{l \neq 1, l \neq j} v_l\right)^2,$$

which is easily seen to be true.

For the second equality in (4.11) we have to show that

$$\left(-\sqrt{v_1 v_2} \prod_{l=3}^{k} v_l\right)\left(\sum_{i=1}^{k} \prod_{l \neq i} v_l\right) = \left(\sum_{j \neq 1} \prod_{l \neq j} v_l\right)\left(-\sqrt{v_1 v_2} \prod_{l=3}^{k} v_l\right)$$

$$+ \left(\sum_{j \neq 2} \prod_{l \neq j} v_l\right)\left(-\sqrt{v_1 v_2} \prod_{l=3}^{k} v_l\right)$$

$$+ \sum_{i=3}^{k} v_i \sqrt{v_1 v_2} \prod_{l \neq 1, l \neq i} v_l \prod_{l \neq 2, l \neq i} v_l.$$

Dividing both sides by $-\sqrt{v_1 v_2} \prod_{l=3}^{k} v_l$ yields

$$\sum_{i=1}^{k} \prod_{l \neq i} v_l = \sum_{j \neq 1} \prod_{l \neq j} v_l + \sum_{j \neq 2} \prod_{l \neq j} v_l - \sum_{i=3}^{k} \prod_{l \neq i} v_l,$$

which is obviously true. This establishes the idempotency of $D$.

For the second statement in the lemma, note that the rank of an idempotent matrix is equal to its trace. It is easily seen that the trace of $D$ is $k - 1$.

<div align="right">□</div>

## REFERENCES

ALY, E.-E. A. A. (1986). Quantile-quantile plots under random censorship. *J. Statist. Plann. Inference* **15** 123–128.

BEIRLANT, J. and DEHEUVELS, P. (1990). On the approximation of P-P and Q-Q plot processes by Brownian bridges. *Statist. Probab. Lett.* **9** 241–251.

DEHEUVELS, P. and EINMAHL, J. H. J. (1992). Approximations and two-sample tests based on P-P and Q-Q plots of the Kaplan–Meier estimators of lifetime distributions. *J. Multivariate Anal.* **43** 200–217.

DOKSUM, K. A. (1974). Empirical probability plots and statistical inference for nonlinear models in the two-sample case. *Ann. Statist.* **2** 267–277.

DOKSUM, K. A. (1977). Some graphical methods in statistics. A review and some extensions. *Statist. Neerlandica* **31** 53–68.

DOKSUM, K. A. and SIEVERS, G. L. (1976). Plotting with confidence: graphical comparisons of two populations. *Biometrika* **63** 421–434.

FISHER, N. I. (1983). Graphical methods in nonparametric statistics: a review and annotated bibliography. *Internat. Statist. Rev.* **51** 25–58.

FLEMING, T. R. and HARRINGTON, D. P. (1991). *Counting Processes and Survival Analysis*. Wiley, New York.

HOLLANDER, M., MCKEAGUE, I. W. and YANG, J. (1997). Likelihood ratio-based confidence bands for survival functions. *J. Amer. Statist. Assoc.* **92** 215–226.

LI, G. (1995). On nonparametric likelihood ratio estimation of survival probabilities for censored data. *Statist. Probab. Lett.* **25** 95–104.

LI, G., HOLLANDER, M., MCKEAGUE, I. W. and YANG, J. (1995). Nonparametric likelihood ratio confidence bands for quantile functions from incomplete survival data. *Ann. Statist.* **23** 628–640.

NAIK-NIMBALKAR, U. V. and RAJARSHI, M. B. (1997). Empirical likelihood ratio test for equality of $k$ medians in censored data. *Statist. Probab. Lett.* **34** 267–273.

NAIR, V. N. (1978). *Graphical Comparisons of Populations in Some Non-linear Models*. Ph.D. dissertation, Univ. California, Berkeley.

NAIR, V. N. (1982). Q-Q plots with confidence bands for comparing several populations. *Scand. J. Statist.* **9** 193–200.

OWEN, A. (1988). Empirical likelihood ratio confidence intervals for a single functional. *Biometrika* **75** 237–249.

OWEN, A. (1990). Empirical likelihood ratio confidence regions. *Ann. Statist.* **18** 90–120.

PAULAUSKAS, V. and RAČKAUSKAS, A. (1989). *Approximation Theory in the Central Limit Theorem. Exact Results in Banach Spaces*. Kluwer, Dordrecht.

PRESS, W. H., TEUKOLSKY, S. A., VETTERLING, W. T. and FLANNERY, B. P. (1992). *Numerical Recipes in C*, 2nd ed. Cambridge Univ. Press.

SHORACK, G. R. and WELLNER, J. A. (1986). *Empirical Processes with Applications to Statistics*. Wiley, New York.

SWITZER, P. (1976). Confidence procedures for two-sample problems. *Biometrika* **63** 13–25.

THOMAS, D. R. and GRUNKEMEIER, G. L. (1975). Confidence interval estimation of survival probabilities for censored data. *J. Amer. Statist. Assoc.* **70** 865–871.

DEPARTMENT OF MATHEMATICS
  AND COMPUTING SCIENCE
EINDHOVEN UNIVERSITY OF TECHNOLOGY
P.O. BOX 513
5600 MB EINDHOVEN
THE NETHERLANDS
E-MAIL: einmahl@win.tue.nl

DEPARTMENT OF STATISTICS
FLORIDA STATE UNIVERSITY
TALLAHASSEE, FLORIDA 32306-4330
E-MAIL: mckeague@stat.fsu.edu