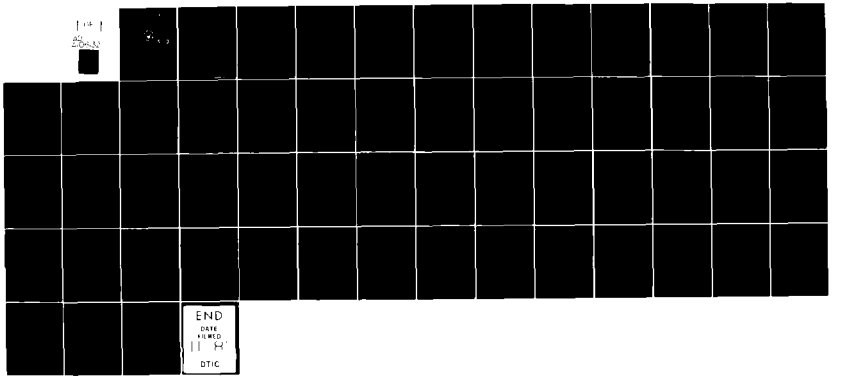
AD-A106 132 NAVAL POSTGRADUATE SCHOOL MONTEREY CA F/G 12/1
QUANTILE ESTIMATION IN DEPENDENT SEQUENCES.(U)
SEP 81 P HEIDELBERGER, P A LEWIS
UNCLASSIFIED NPS55-81-015 NL

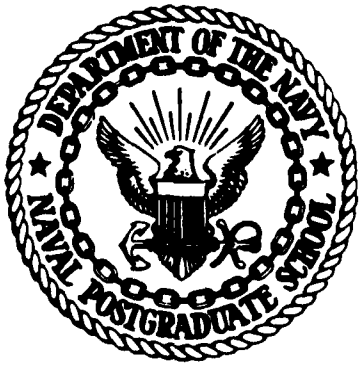


(1) LEVEL II

NPS55-81-015

NAVAL POSTGRADUATE SCHOOL

Monterey, California



QUANTILE ESTIMATION IN DEPENDENT SEQUENCES

by

P. Heidelberger

P. A. W. Lewis

September 1981

Approved for public release; distribution unlimited

Prepared for:
Chief of Naval Research
Arlington, VA 22217

NAVAL POSTGRADUATE SCHOOL
MONTEREY, CALIFORNIA

Rear Admiral J. J. Ekelund
Superintendent

David A. Schrady
Acting Provost

This work was supported in part by the Office of Naval Research under Grant NR-42-284.

P. Heidelberger (by P.A.W.L.)
P. Heidelberger
IBM Research Center

P. A. W. Lewis
P. A. W. Lewis, Professor
Department of Operations Research

Reviewed by:

K. T. Marshall
K. T. Marshall, Chairman
Department of Operations Research

Released by:

William M. Tolles
William M. Tolles
Dean of Research

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

REPORT DOCUMENTATION PAGE	READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER: NPS55-81-015	2. GOVT ACCESSION NO.: AD-A106132
	3. RECIPIENT'S CATALOG NUMBER:
4. TITLE *(and Subtitle)*: QUANTILE ESTIMATION IN DEPENDENT SEQUENCES	5. TYPE OF REPORT & PERIOD COVERED: Technical Report
	6. PERFORMING ORG. REPORT NUMBER:
7. AUTHOR(s): P. Heidelberger; P. A. W. Lewis	8. CONTRACT OR GRANT NUMBER(s):
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Naval Postgraduate School, Monterey, CA 93940	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 61153N; RR014-05-01; N0001481WR10001
11. CONTROLLING OFFICE NAME AND ADDRESS: Chief of Naval Research, Arlington, VA 22217	12. REPORT DATE: September 1981
	13. NUMBER OF PAGES: 53
14. MONITORING AGENCY NAME & ADDRESS *(if different from Controlling Office)*	15. SECURITY CLASS. *(of this report)*: UNCLASSIFIED
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT *(of this Report)*

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT *(of the abstract entered in Block 20, if different from Report)*

18. SUPPLEMENTARY NOTES

19. KEY WORDS *(Continue on reverse side if necessary and identify by block number)*

Quantile; quantile estimation; dependent sequences; nonparametric quantile estimators; ϕ-mixing process; maximum transformation; extreme quantile; confidence interval estimates; averaged group quantiles; nested group quantiles.

20. ABSTRACT *(Continue on reverse side if necessary and identify by block number)*

Standard nonparametric estimators of quantiles based on order statistics can be used not only when the data are i.i.d., but also when the data are from a stationary, ϕ-mixing process of continuous random variables. However, when the random variables are highly positively correlated, sample sizes needed for acceptable precision in estimates of extreme quantiles are computationally unmanageable. A practical scheme is given, based on a maximum transformation in a two-way layout of the data, which reduces the

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

sample size sufficiently to allow an experimenter to obtain a point estimate of an extreme quantile. Three schemes are then given which lead to confidence interval estimates for the quantile. One uses a spectral analysis of the reduced sample. The other two, averaged group quantiles and nested group quantiles, are extensions of the method of batched means to quantile estimation. None of the schemes requires that the process being simulated is regenerative.

QUANTILE ESTIMATION IN DEPENDENT SEQUENCES

by

P. Heidelberger
IBM Research Center
Yorktown Heights, NY 10598

P. A. W. Lewis
Naval Postgraduate School
Monterey, CA 93940

ABSTRACT

Standard nonparametric estimators of quantiles based on order statistics can be used not only when the data are i.i.d., but also when the data are from a stationary, ϕ-mixing process of continuous random variables. However, when the random variables are highly positively correlated, sample sizes needed for acceptable precision in estimates of extreme quantiles are computationally unmanageable. A practical scheme is given, based on a maximum transformation in a two-way layout of the data, which reduces the sample size sufficiently to allow an experimenter to obtain a point estimate of an extreme quantile. Three schemes are then given which lead to confidence interval estimates for the quantile. One uses a spectral analysis of the reduced sample. The other two, averaged group quantiles and nested group quantiles, are extensions of the method of batched means to quantile estimation. None of the schemes requires that the process being simulated is regenerative.

QUANTILE ESTIMATION IN DEPENDENT SEQUENCES

1. INTRODUCTION

Let $\{X_n\}$ be a sequence of random variables which is assumed to be stationary with marginal distribution $F_X(x) = P\{X_n \leq x\}$. For any p, $0 < p < 1$, the p-quantile of $F_X(x)$ satisfies the equation $p = F_X(x_p)$ if $F_X(x)$ has a positive density in a neighborhood of x_p. This paper is concerned with the estimation of quantiles x_p from samples X_n, $n=1,\ldots,N$. In particular the paper is concerned with both point estimates and confidence interval estimates for quantiles when the X_n's are highly correlated. This is the usual case when the X_n's are the outputs from a simulation of a stochastic system, for example the waiting times in a queue.

In the case of independent X_n's, methods for quantile estimation are well known (e.g. Connover, 1980, pp. 71). Thus let $X_{(n)}$ denote the order statistics from the sample of size N, and let $\lfloor x \rfloor$ denote the greatest integer less than or equal to x. Then a standard nonparametric estimator

$$\hat{x}_p = X_{(\lfloor pN + 1 \rfloor)}$$

is a consistent, asymptotically normal, estimator of x_p with bias which is $O(N^{-1})$ and variance $[p(1-p)]/[f_X^2(x_p)N] + O(N^{-2})$, where $f_X(x_p)$ is the derivative of $F_X(x)$ at x_p. This variance can be estimated by estimating $f_X(x_p)$, but nonparametric confidence intervals can also be obtained for x_p using the order statistics of the sample (Connover, 1980, pp. 111-116).

The difficulty with these estimation methods is that they require a large amount of computer storage and computing time to sort the sample. Also for high or low p, the order-statistic estimator might be biased and very non-normal. A solution is to use the maximum transformation (Goodman, Lewis and Robbins, 1971). This not only moves the problem back to one of estimating a median, but also allows the use of stochastic approximation (Robbins-Monro) methods. A very satisfactory method of this type using a minimal amount of memory and giving both point and confidence interval estimates for the unknown quantile x_p has been given by Robinson (1975).

For the case of dependent X_n's, the usual situation encountered in system simulation studies, quantile estimation is much more difficult than in the independent case. The point estimate $\hat{x}_p$ is still valid; however its variance is inflated by a factor $p(0;x_p)$. Here $p(0;x_p)$ is the initial point on the spectrum of the binary process $\{I_n(x_p)\}$, defined for each n to be 1 if $X_n \leq x_p$ and zero otherwise. More directly it is

$$p(0;x_p) = \lim_{n\to\infty} n \, \mathrm{Var}\{(I_1(x_p) + \ldots + I_n(x_p))/n\} .$$

The nonparametric confidence interval estimation procedures are no longer directly applicable. However it is possible to estimate $p(0;x_p)$ using methods of Heidelberger and Welch (1980, 1981) and to estimate the density $f(x_p)$ by standard methods; used together these give an estimate of the variance of $\hat{x}_p$

and large-sample confidence intervals. The main problem with these order-statistic estimates is, however, that the sample sizes required to obtain reasonable precision with the estimates are prohibitive when the X_n's are highly positively correlated.

The basic scheme considered in this paper to handle the sample size problem for dependent, ϕ-mixing sequences is to set out the data in a (conceptual) $v \times m$ array

$$X_{k,i} = X_i + (k-1)m \ ,$$

for $k = 1,\ldots,v$ and $i = 1,\ldots m = N/v$, where v is often but not necessarily chosen so that $p^v = 0.5$. Then if m(and N) is large enough so that X_n's which are m apart are approximately independent, the table can be collapsed by taking maxima down the columns. The maxima

$$Y_i = \max_{1 \le k \le v} X_{k,i} \ ,$$

for $i = 1,\ldots,m$ are now a reduced sample whose $q=p^v$ quantile, y_q , corresponds to the desired quantile, x_p , of the $\{X_n\}$ process. The main point of this procedure is that it gives a very substantial sample size reduction. In addition, there is generally slightly less overall correlation in the Y_i sequence and this counteracts the increase in variance which occurs with the use of the maximum transformation and order statistic estimates of the quantile x_p .

Since the Y_i's (the maximum transformed sample) are still correlated several problems remain, notably estimating the variance of the point estimate $\hat{y}_q$ (the $q=p^v$ sample quantile of the Y_i's) and obtaining confidence interval estimates of x_p. Three confidence interval schemes are considered; averaged group quantiles, nested group quantiles and a scheme based on estimating the probability density function of the Y_i's at x_p, and the initial spectral point, $p(0;\hat{y}_q)$ of the binary process obtained from comparison of the Y_i's to $\hat{y}_q$, i.e. $I_i(\hat{y}_q)$.

Extensive empirical sampling studies using M/G/1 queues and stationary sequences of autocorrelated exponential random variables show that the above schemes provide a reliable method for estimating a quantile in a dependent sequence $\{X_n\}$

2. QUANTILE ESTIMATION AND THE MAXIMUM TRANSFORMATION

1. Quantile estimation for i.i.d. sequences.

Let $X_1,\ldots,X_n,\ldots,X_N$ be a sample of i.i.d. random variables from a continuous distribution $F_X(x)$ with probability density function $F_X(x)$. For $0 < p < 1$ let

$$x_p = \inf\{x : F_X(x) = p\} = F_X^{-1}(p) \quad , \tag{2.1}$$

where $F_X^{-1}(p)$ is the inverse of $F_X(x)$ with derivative $1/f_X(x_p)$. The quantity x_p is the pth quantile of $F_X(x)$.

Let $X_{(1)} \leq \ldots \leq X_{(n)} \leq \ldots X_{(N)}$ be the order statistics corresponding to the sample. The usual nonparametric point estimator of x_p is the pth sample quantile

$$\hat{x}_p = X_{(\lfloor Np+1 \rfloor)}, \qquad 0 < p < 1 \tag{2.2}$$

where $\lfloor z \rfloor$ denotes the integral part of z. The following properties of $\hat{x}_p$ are well known (David, 1970, 65-67):

$$E(\hat{x}_p) = x_p - \frac{f'_X(x_p)}{2f^3_X(x_p)} \cdot \frac{p(1-p)}{N+2} + O\left(\frac{1}{N^2}\right); \tag{2.3}$$

$$\text{var}(\hat{x}_p) = \sigma_p^2 = \frac{p(1-p)}{Nf^2_X(x_p)} + O\left(\frac{1}{N^2}\right) \tag{2.4}$$

and

$$N^{\frac{1}{2}}(\hat{x}_p - x_p) \xrightarrow{D} N(0, N\sigma_p^2) \tag{2.5}$$

as $N \to \infty$.

Nonparametric confidence intervals for x_p are based on the identity (see e.g. Connover, 1980, pp. 111-116)

$$P\{X_{(L)} \leq x_p < X_{(U)}\} = \sum_{j=L}^{U-1} \binom{N}{j} p^j (1-p)^{N-j}, \tag{2.6}$$

with $L < U$ chosen in such a way as to make the probability as close as possible to the desired coverage $(1-\alpha)$.

Another method for estimating quantiles uses the method of stochastic approximation (Robbins and Monro (1981)). This provides a method of quantile estimation which avoids sorting and requires only a minimal amount of storage. The asymptotic variance of the stochastic approximation estimate is given by (2.4). A very satisfactory development of this method is given in Robinson (1975). Since it is not extendable to dependent data, it is not described. The key idea in the method, however, is the use of the maximum transformation, and since this is of use in the present context, it is described next.

2.2. The maximum transformation

The maximum (minimum) transformation is a computationally simple transformation and compaction of i.i.d. data which transforms the problem of estimating an extreme quantile of a random variable into that of estimating a more central quantile, e.g. the median. Thus assume that $p > 1/2$, and let $Y = \max(X_1,\ldots,X_v)$. (When $p < 1/2$ a minimum transformation is used). Then

$$P\{Y \leq x_p\} = F_Y(x_p) = \{F_X(x_p)\}^v = p^v \equiv q \quad . \tag{2.7}$$

Thus the pth quantile of F_X is the $q = p^v$th quantile of the Y variable. In particular let $v = \lfloor \ell n(1/2)/\ell n(p) \rfloor$; then x_p is approximately the median of the Y variable. Details are given in Goodman, Lewis, and Robbins (1971).

There is a price paid with this transformation in inflation of the leading term of the variance (2.4). Thus let v

divide N, so that $N/v = m$. Then the X sample is reduced to a sample $Y_1,\ldots,Y_m$ by application of the maximum transformation to each successive non-overlapping group of X_i's of size v. Then the density of the Y_i's is $f_y(x_p) = v\, f_X(x_p)p^{v-1}$ and the variance of the order statistic estimator, $\hat{y}_q$, of x_p in the sample $Y_1,\ldots,Y_m$ is

$$\text{(2.8)} \qquad \operatorname{var}(\hat{y}_q) \sim \frac{p(1-p)}{Nf_X^2(x_p)} \cdot \frac{(1-p^v)}{v(1-p)p^{v-1}} \sim \operatorname{var}(\hat{x}_p)\, \frac{(1-p^v)}{v(1-p)p^{v-1}} .$$

The right-hand multiplicative factor in this last expression is approximately 1.4 if v is chosen to make $q = p^v$ approximately 0.5 .

When using stochastic approximation, shifting to the neighborhood of the median is essential for the method to be a well-behaved estimation procedure. For estimating several quantiles simultaneously, the method is applicable in a nested scheme which is very efficient with respect to storage and speed.

For $p < 1/2$ a minimum transformation is used. Next-to-maximum and maximum-minimum schemes give more flexible and robust schemes but are not discussed in the present paper.

2.3. Dependent Data: General Considerations and Order Statistics

Quantile estimation in dependent data is an order of magnitude more difficult than in independent data. Fishman (1978, p. 270) notes that no satisfactory solution exists. Iglehart (1976), Seila (1976) and Moore (1979) have given special

methods for processes $\{X_n\}$ with regenerative structure, i.e. processes for which there exist random time points at which the process restarts probabilistically. An example is the waiting time process $\{W_n\}$ in the M/G/1 queue which regenerates every time a customer arrives to find the queue empty, so that the waiting time of that customer is zero.

It is possible to use the order statistic estimator $\hat{x}_p$ given at (2.2), where we ignore for now the problem of the initial transient which is usually encountered in simulations. Conditions for convergence and Central Limit Theorems for sample quantiles are well known (Sen, 1972; Babu and Singh, 1978). Thus let

$$I_n(x) = 1 \quad \text{if} \quad X_n \leq x \,, \qquad = 0 \quad \text{otherwise} \,; \tag{2.9}$$

$$\bar{I}(x) = \sum_{n=1}^{N} I_n(x)/N \,, \tag{2.10}$$

the empirical c.d.f. at x, and

$$p(0;x) = \lim_{N\to\infty} \{N \operatorname{var} \bar{I}(x)\} = \sum_{k=-\infty}^{\infty} \gamma_k(x) \,, \tag{2.11}$$

where

$$\gamma_k(x) = \operatorname{cov}\{I_n(x), I_{n+k}(x)\} \qquad k = 0,\pm 1,\ldots \quad -\infty < x < +\infty \tag{2.12}$$

$$= P\{X_n \leq x \ , \ X_{n+k} \leq x\} - P\{X_n \leq x\} \ P\{X_{n+k} \leq x\} \ .$$

The notation $p(0;x_p)$ comes from the fact that this quantity is the initial point (f=0) of the spectrum of the process $\{I_n(x_p)\}$;

$$(2.13) \qquad p(f;x) = \sum_{k=-\infty}^{\infty} \cos(2\pi fk) \ \gamma_k(x) \ , \qquad -\frac{1}{2} \leq f \leq 1/2$$

Now the Central Limit Theorem for $\hat{x}_p = X_{(\lfloor np+1 \rfloor)}$ (Sen, 1972) states that if $f_X(x)$ is continuous, bounded and non-zero in some small neighborhood of x , and if $f_X'(x)$ is bounded in this neighborhood, then

$$(2.14) \qquad \frac{N^{1/2}(\hat{x}_p - x_p)}{(p(0;x_p))^{1/2}/f_X(x_p)} \sim N(0,1)$$

if the process $\{X_n\}$ is ϕ-mixing and

$$(2.15) \qquad \sigma^2_{x_p} = p(0;x_p)/\{f_X(x_p)\}^2$$

is finite. For details on ϕ-mixing see Sen (1972) and Billingsley (1968); some discussion is given in Section 2.4.

Regenerative processes such as the M/G/1 queuing system waiting time $\{W_n\}$ are ϕ-mixing.

The problem with using $\hat{x}_p$ as an estimator in positively correlated sequences is that the sample sizes required for adequate precision are prohibitively large. Both sorting times

and memory times are then unrealistic. A measure of the inflation of sample size over the independence case is the ratio of $p(0;x_p)$ to its value $p(1-p)$ for the i.i.d. process with identical marginal distributions. This has been investigated by Blomqvist (1967) for the M/G/1 queue. For extreme quantiles of the M/M/1 queue with traffic intensity $\rho = 0.9$ this ratio is 400 . Greater ratios are possible, depending on the traffic intensity and the skewness of the service time distribution. Specifically, the M/M/1 queue with $\rho = 0.9$ requires a sample size of roughly 500,000 customers to estimate the 0.99 quantile of the waiting time distribution to within plus or minus 10% accuracy. This is derived from Table 8 of Blomqvist (1967); more precisely this is the sample size required for a 90% confidence interval for $x_{.99}$ to have a relative half-width of 0.10 . For the 0.999 quantile the required sample size is approximately 2,300,000. Clearly storing and sorting the entire sequence is impractical in such cases. Actually to produce an order-statistic point estimate of x_p requires storing only the largest $(1-p)N$ values of the sequence. However, this ordering must be dynamically maintained as the sequence is generated, a computationally expensive operation. Additional storage is required to estimate the variance term $p(0;x_p)$: the positions in the sequence at which the $(1-p)N$ largest values occur must also be saved in order to construct the binary sequence $\{I_n(\hat{x}_p)\}$. Furthermore, the Discrete Fourier Transform of the extremely long sequence $\{I_n(\hat{x}_p)$, $n = 1,\ldots,N\}$ must be taken to accomplish the variance estimation.

Some of the storage and sorting problems could be relieved by using stochastic approximation . However properties of the stochastic approximation, particularly its asymptotic variance, are unknown for dependent data. More importantly direct application of the maximum transformation to bring the quantile estimation down to a problem of estimating the median requires independence of the X_n's .

We now present a method for using the maximum transformation t. achieve sample size reduction and a practical scheme for point and confidence interval estimates with dependent data.

2.4. The maximum transformation in the dependent case

The basic idea behind the use of the maximum transformation is to combine elements of X_n , $n = 1,2,...,$ in a ϕ-mixing process which are sufficiently far apart so as to be approximately independent. To define ϕ-mixing let $M_{-\infty}^{n}$ and M_{n+m}^{∞} be respectively the σ-fields generated by $\{X_i;\ i \leq n\}$ and $\{X_i;\ i \geq n + m\}$. If $E_1 \epsilon M_{-\infty}^{n}$ and $E_2 \epsilon M_{n+m}^{\infty}$, then $\{X_n\}$ is Q-mixing if for all $n(-\infty < n < \infty)$ and $m(\geq 1)$

$$(2.16) \qquad |P(E_2|E_1) - P(E_2)| \leq \phi(m) , \qquad \phi(m) > 0 ,$$

where $1 \geq \phi(1) \geq \phi(2) \geq \ldots$ and $\lim_{m\to\infty} \phi(m) = 0$.

Thus we set out the data in a $v \times m$ array, where $m = N/v$, as

(2.17) $\quad X_{k,i} = X_{i+(k-1)m} \qquad k=1,\ldots,v;\ i=1,\ldots,m\ .$

We assume for the moment that elements in the m columns are independent; since we are assuming that v is fixed and the process $\{X_n\}$ is ϕ-mixing this will certainly be true as N (and m) get very large. Now applying the maximum transform down the columns we get a reduced series Y_i, where

(2.18) $\quad Y_i = \max_{1\le k\le v} X_{k,i} \qquad i = 1,2,\ldots,m\ .$

Values of v required to reduce the problem to that of estimating the median of the Y_i's are given in Table 2.1 .

Table 2.1

Size of v for $p^v \doteq 1/2$

p	0.900	0.950	0.975	0.980	0.990	0.995	0.998	0.999
v	6.6	13.5	27.4	34.3	69.0	138.3	346.2	692.8

We emphasize here that transformations to points other than the median are possible, and are essential when considering simultaneous estimation of several quantiles or estimation of the median of the X_n's . However in what follows transformation to the median is usually assumed. There will be an attendant bias because p^v will never quite be $1/2$.

The purposes of the maximum transformation are

(i) to reduce the attendant sample size. It should be kept in mind

that for higher quantiles larger sample sizes are needed for given precision. The result is that for fixed precision the Y_i series is roughly the same length for all $p > 0$. In particular for $p = 0.999$ the reduction of the sample size is by a factor of 693 which is sufficient to reduce a series of completely unmanageable length to one which can clearly be accommodated in a digital computer (see Section 5).

(ii) To reduce the problem to estimation of a median rather than an extreme quantile. Since stochastic approximation is not used with the Y_i series, this is not essential for point estimates. It is, however, helpful in obtaining confidence interval estimates.

(iii) To possibly reduce correlation. It will be shown that the correlations in the $\{Y_i\}$ series are slightly less than the correlations in the $\{X_n\}$ series. Though a small effect, it is usually sufficient to offset inflation in variances due to use of the maximum transformation. To see this assume that the structure of $I_n(x_p)$ is Markovian with $p(0;x_p) \simeq (1+\rho)/(1-\rho)$. If $\rho = 0.99$ this equals 199. If ρ for the Y_i process chopped at its median, i.e. $\{I_n(x_{0.5})\}$, is reduced to only 0.98, then the new $p(0;x_p) \simeq 99$.

Clearly if m is too small, the assumption of independence down the column will be violated. Some analysis of this effect is possible. Expanding the definition (2.18) to account explicitly for the parameter v, the number of X_n's, v apart, over which the maximum is taken, so that $Y_i(v) = \max_{1 \le k \le v} X_{k,i}$, $i=1,2,\ldots,m$, we have from Billingsley (1968, pp. 174, 20.49) that

(2.19) $$|P\{Y_i(v) \leq x_p\} - p^v\}| \leq v\phi(m) \ .$$

Thus we want m as large as possible and, for given v, the difference in the probabilities goes to zero as $m \to \infty$, by the definition of mixing. A slightly tighter bound is given by the

Lemma. $$|P\{Y_i(v) \leq x_p\} - p^v\} \leq \phi(m) \sum_{k=0}^{v-2} p^k$$

for $v \geq 2$.

A straight forward proof of this is obtained using induction.

The point estimate used for x_p is

$$\hat{y}_q = Y_{(\lfloor mq+1 \rfloor+1)}$$

where $q = p^v$. Since the Y_i's are dependent the properties of this estimator are given as at (2.14) and (2.15), with proviso for some peculiarities in the structure of the $\{Y_i\}$ process which we discuss now. Finite sample properties of the estimator $\hat{y}_q$ are studied by simulation in Section 5.

2.5. The maximum-transformed process $\{Y_i\}$.

The maximum-transformed process $\{Y_i\}$, obtained from the $\{X_n\}$ process by taking the maximum of X_n's which are m apart, is not a stationary process in the same sense that the $\{X_n\}$ process is stationary. To see this consider

$$(2.20) \qquad Y_1 = \max_{1 \le k \le v} X_{k,1} = \max\{X_1, X_{m+1}, \ldots, X_{(v-1)m+1}\} .$$

It will be correlated with successive Y_i's , and this correlation can be expected to decrease since the $\{X_n\}$ process is mixing. However, the correlation will eventually increase because

$$(2.21) \qquad Y_m = \max\{X_m, X_{2m}, \ldots, X_{vm}\}$$

is coupled to Y_1 by the $v-1$ dependent pairs $\{X_m, X_{m+1}\}, \ldots \{X_{(v-1)m}, X_{(v-1)m+1}\}$. Thus the $\{Y_i\}$ process is circular. However, it is not stationary in a circular sense because the correlation between Y_1 and $Y_2 = \max\{X_2, X_{m+2}, \ldots, X_{(v-1)m+2}\}$ is through the v dependent pairs $\{X_1, X_2\}$, $\{X_{m+1}, X_{m+2}\}, \ldots, \{X_{(v-1)m+1}, X_{(v-1)m+2}\}$, and is therefore different from the correlation between Y_1 and Y_m . In a circularly stationary process these correlations would be the same. Towards the middle of the process, say $i = m/2$, the lag one correlations will be the same and there will be no "edge effect" because, if m is large enough, the correlation between $Y_{m/2}$ and Y_1 , and $Y_{m/2}$ and Y_m will be zero.

Without belabouring this "edge effect" and circularity in the $\{Y_i\}$ process there are two main points to be made.

(i) For finite m it will introduce bias because, for instance, methods using spectral techniques for estimating $p(0;x)$ in Section 3 are based on assumptions of stationary.

Note, however, that $\{Y_i\}$ is marginally stationary, so density estimation techniques for $f_y(x)$ are not affected.

(ii) For m large the effect will be negligible and the asymptotics go through in the usual way by ignoring a fixed (or slowly increasing) set of Y_i's at the beginning and end of the process. The usual assumption is that the extent of this set is smaller than the extent of the dependence in the $\{Y_i\}$ process.

3. CONFIDENCE INTERVALS FOR QUANTILES IN DEPENDENT DATA

In the previous section we gave a point estimator $\hat{y}_q$ for the quantile of the marginal distribution in a stationary time series. In this section we consider three methods for generating confidence interval estimates for these quantiles. The first method is an extension of the spectral method of Heidelberger and Welch (1980; 1981). The second and third methods are extensions of the method of batch means (see e.g. Law and Carson (1979) or Fishman (1978)). These last two methods also give two new point estimates.

These methods for generating confidence interval estimates do not rely upon use of the max-transform and the case $v = 1$ corresponds to working with the original time series. Moreover, if the max-transform is used we do not require that $p^v = 0.50$, i.e. $v \sim \ell n.5/\ell n\ p$. There are considerable computational savings, however, from using some maximum transformation. This and statistical considerations will generally dictate a maximum transformation using a v slightly less than the median transform v. Another determing factor is that one generally wants to estimate more than one quantile. This factor will be considered elsewhere.

3.1. A Spectral Method

Let $\{Y_1,\ldots,Y_m\}$ be the sequence of max-transformed variables as defined by equation (2.18). The point estimate of x_p is the qth order statistic of $\{Y_1,\ldots,Y_m\}$, $\hat{y}_q$, where $q = p^v$. Let $f_y(x)$ be the stationary density function of Y_i

and let $p(0;x)$ be the stationary spectral density of the indicator function process $I_i(x)$ associated with $\{Y_i\}$ at zero frequency as defined by equation (2.13). Here edge effects because of the quasi-circularity of the Y_i process are ignored. For large values of m (see Sen (1972) and Babu and Singh (1978)) $\sqrt{m}\ (\hat{y}_q - x_p)/(p(0;x_p)/f_y^2(x_p))^{1/2}$ has a normal distribution with mean zero and variance one. Confidence intervals for x_p based on this Central Limit Theorem are generated by estimating both $p(0;x_p)$ and $f_y(x_p)$ at the estimated quantile . This is the first method alluded to above.

The quantity $p(0;x_p)$ is estimated using the spectral method of Heidelberger and Welch (1980; 1981) applied to the sequence $\{I_i(\hat{y}_q)\ ,\ i=1,\ldots,m\}$. This method uses least squares to fit a low order polynomial of degree d to the logarithm of the first K values of the averaged periodogram of $\{I_i(\hat{y}_q)\}$. As suggested in Heidelberger and Welch (1980; 1981) we used $K = 25$ and $d = 2$, although for extreme quantiles in highly congested queues with short run lengths, $d = 3$ was required to produce valid confidence intervals. This point is discussed more fully in Section 4.

We found little or no loss in confidence interval coverage by using the estimated quantile $\hat{y}_q$ with the sequence $\{I_i(\hat{y}_q)\}$ rather than using the known quantile x_p with the sequence $\{I_i(x_p)\}$. However, a proof of the convergence of the distributional properties of the periodogram of $\{I_i(\hat{y}_q)\}$ appears difficult.

The density $f_y(x_p)$ of the max-transformed variables $Y_1,\ldots,Y_m$ may be estimated using a standard kernel density estimator (see Parzen (1962) and Rosenblatt (1956)). Specifically, let $W(x)$ be a weighting function such that

$$0 \leq W(x) \leq W < \infty ,$$

$$\int_{-\infty}^{\infty} W(x)dx = 1 , \tag{3.1}$$

$$\lim_{|x|\to\infty} |xW(x)| = 0 .$$

Let $b(m)$ be a sequence of bandwidth constants satisfying $\lim_{m\to\infty} b(m) = 0$ and $\lim_{m\to\infty} mb(m) = \infty$. For any x the kernel density estimate of $f_y(x)$ is defined by

$$\hat{f}_y(x) = \frac{1}{m} \sum_{j=1}^{m} \frac{1}{b(m)} W((x-Y_j)/b(m)) . \tag{3.2}$$

Under the conditions (3.1) and if $\sum_{n=0}^{\infty} \phi(n)^{1/2} < \infty$, where the $\phi(n)$'s are defined at (2.16), then $\hat{f}_y(x)$ converges in probability to $f_y(x)$.

In our case we require a density estimate at the <u>unknown point</u> x_p. Simple Taylor series expansions show that $\hat{f}_y(\hat{y}_q)$ converges in probability to $f_y(x_p)$ if the first k derivatives, $W^{(k)}(x)$, of $W(x)$ satisfy $|W^{(k)}(x)| \leq W < \infty$, if $\int_{-\infty}^{\infty} |W^{(k)}(x)|dx < \infty$ and $\lim_{|x|\to\infty} |xW^{(k)}(x)| = 0$ and if

$b(m) = Z_m m^{-c}$ where Z_m converges in probability to a positive constant and $0 < c < k/2(k+1)$. If the weighting function $W(x)$ does not satisfy the different ability conditions then more delicate arguments are required to show convergence of $\hat{f}_y(\hat{x}_p)$ (see, for example, Robinson (1975)).

We have experimented with two weighting functions, a triangular window

$$W(x) = \begin{cases} 1 - |x| & |x| \leq 1 , \\ 0 & |x| > 1 , \end{cases} \tag{3.3}$$

and a cosine window

$$W(x) = \begin{cases} 1/2 \cos(x) & |x| \leq \pi/2 , \\ 0 & |x| > \pi/2 ; \end{cases} \tag{3.4}$$

Bandwidth sequences $b(m)$ had the form $b(m) = Z_m m^{-c}$ for various powers of $0 < c < 1/2$ and random variables Z_m which ranged from constants to measures of the spread of the distribution of the Y_j's . We found the density estimates to be relatively insensitive to both the shape of the weighting function and the parameters of the bandwidth sequence. For small samples they tended to slightly overestimate $f_y(x_p)$ but converged to $f_y(x_p)$ for large samples. In Section 4 we report the results of experiments using the triangular window with $c = 1/3$ and Z_m equal to the inter-quantile range of the Y_j's , i.e. $b(m) = \left\{ Y_{(\lfloor .75m+1 \rfloor)} - Y_{(\lfloor .25m+1 \rfloor)} \right\} m^{-1/3}$.

3.2. Averaged and Nested Group Quantiles: Definition

We now describe two methods of generating confidence intervals for the quantile x_p which do not require direct estimation of $p(0;x_p)$ for the $\{Y_i\}$ sequence. These methods, which we call averaged group quantiles (agq) and nested group quantiles (ngq), are extensions of the method of batch means to quantile estimation. Seila (1976) considered a version of agq without the max transform (v=1) for regenerative processes.

Let N be the total sample size and divide $\{X_1,\ldots,X_N\}$ into G non-overlapping groups with mv observations in each group $(N = Gmv)$. Define

$$X_{\ell,j,k} = X_{(\ell-1)mk+j+(k-1)m} \quad \text{for} \quad \begin{array}{l}\ell=1,\ldots,G,\\ j=1,\ldots,m,\\ k=1,\ldots,v.\end{array} \tag{3.5}$$

The subscript ℓ refers to the group number, and the data in each group can be thought of as forming a matrix of dimension v by m. The first row of this matrix is formed from the first m observations in the group, the second row is made up of the second m observations in the group, etc. For any fixed values of ℓ and j the points $X_{\ell,j,k}$ and $X_{\ell,j,k+1}$ $(1\le k<v)$ are at lag m (i.e., separated by $m-1$ observations) and for large values of m, $X_{\ell,j,1}$, $X_{\ell,j,2},\ldots,X_{\ell,j,k}$ are approximately independent. Define $Y_{\ell j} = \max_{1\le k\le v} X_{\ell,j,k}$; thus $Y_{\ell j}$ is the jth max-transformed variable in the ℓth group. For large m $P\{Y_{\ell j} \le x_p\} \doteq q = p^v$ with $|P\{Y_{\ell j} \le x_p\} - q| \le \phi(m)v$ as in Section 2.

For each $\ell=1,\ldots,G$ let $Y_{\ell(1)} \leq Y_{\ell(2)} \cdots \leq Y_{\ell(m)}$ be the order statistics of $\{Y_{\ell 1}, Y_{\ell 2}, \cdots Y_{\ell m}\}$ and define $\hat{y}_{q\ell} = Y_{\ell(\lfloor mq+1 \rfloor)}$, i.e. $\hat{y}_{q\ell}$ is an order statistic estimate of x_p derived from the ℓth group. We call $\hat{y}_{q\ell}$ a group quantile estimate. As in the case of batch means, for large values of m, $\{\hat{y}_{q1},\ldots,\hat{y}_{qG}\}$ will be distributed as i.i.d. normals with mean x_p and variance $p(0;x_p)/mf_Y^2(x_p)$. This suggests several point and interval estimates for x_p.

The <u>averaged group quantile estimate</u> is defined by

$$\text{agq}(x_p) = \frac{1}{G} \sum_{\ell=1}^{G} \hat{y}_{q\ell} \,. \tag{3.6}$$

A confidence interval for x_p is formed by assuming that

$$\sqrt{G}\,(\text{agq}(x_p) - x_p)/S(\text{agq}) \tag{3.7}$$

has a students-t distribution with $G-1$ degrees of freedom where

$$S^2(\text{agq}) = \frac{1}{G-1} \sum_{i=1}^{G} (\hat{y}_{q\ell} - \text{agq}(x_p))^2 \,.$$

This confidence interval is valid if $\{\hat{y}_{q1},\ldots,\hat{y}_{qG}\}$ are i.i.d. normals with mean x_p and finite variance.

The <u>nested group quantile point estimate</u> of x_p is defined to be the median of $\hat{y}_{q1},\ldots,\hat{y}_{qG}$, i.e. if $\hat{y}_{q(1)} \leq \cdots \leq \hat{y}_{q(G)}$ are the order statistics of $\hat{y}_{q1},\ldots,\hat{y}_{qG}$,

then the ngm is defined by (assuming G is odd)

$$ngq(x_p) = \hat{y}_{q(\lfloor .5G+1 \rfloor)} \cdot \tag{3.8}$$

An approximate $100\times(1-\alpha)\%$ confidence interval for x_p from (2.6) is

$$(\hat{y}_{q(L)} , \hat{y}_{q(U)}) , \tag{3.9}$$

where for any $L < U$, α is defined as at (2.6), by

$$1 - \alpha = \sum_{k=L}^{U-1} \binom{G}{k} (.5)^k (.5)^{G-k} . \tag{3.10}$$

This confidence interval is valid if $\{\hat{y}_{q1}, \ldots, \hat{y}_{qG}\}$ are mutually independent with $P\{\hat{y}_{q\ell} \leq x_p\} = 0.5$. For large samples this is guaranteed by the asymptotic normality of order statistics.

As $m \to \infty$, $agq(x_p)$ and $ngq(x_p)$ are asymptotically unbiased and their confidence intervals are asymptotically valid, since the sequence $\{X_n\}$ is assumed to be mixing. Notice that if $G = 1$, then $agq(x_p) = ngq(x_p) = \hat{y}_q$ where $\hat{y}_q$ is defined by equation (2.20), though clearly $G = 1$ is not reasonable if a confidence interval estimate is required. Furthermore, as with $\hat{y}_q$, these point and interval estimates are valid if the max-transform is not used $(v=1)$ in which case one is working with group quantiles of the original sequence $\{X_k\}$. If the max-transform is used there is no requirement that $p^v = 0.5$.

3.3. Averaged and Nested Group Quantiles: Discussion

There are a number of possible sources of error in the above two confidence interval procedures. First, if m is small then, because of dependence, $P\{Y_{\ell j} \leq x_p\}$ will differ from $q = p^v$ by a significant amount and $E(\hat{y}_{q\ell}) \neq x_p$. Notice that with $G > 1$, m for $ngq(x_p)$ and $agq(x_p)$ is smaller than the m for $\hat{y}_q$ and therefore this source of error is more likely with $ngq(x_p)$ and $agq(x_p)$ than with $\hat{y}_q$. Second, even if the $\hat{y}_{q\ell}$ sequence consists of i.i.d. observations, if m is small $E(\hat{y}_{q\ell}) \neq x_p$ due to the bias in order statistics for small samples. This bias will be more severe at the tails of the distribution. This problem is again potentially more serious with $ngq(x_p)$ and $agq(x_p)$ than with $\hat{y}_q$. Third, the random variables $y_{q1},\ldots,y_{qG}$ may be correlated. All three of these factors suggest that the number of groups, G, should be as small as possible. The choice $G = 5$ is the smallest value of G for which $P\{\hat{y}_{q(1)} \leq x_p < \hat{y}_{q(G)}\} \geq 0.90$ (for $L = 1$, $U = G = 5$ the confidence level is 0.9375). Given that G should be small, any independence tests for $\hat{y}_{q1},\ldots,\hat{y}_{qG}$ will have very low power. The sampling experiments described in Section 4 use $G = 5$ for both $agq(x_p)$ and $ngq(x_p)$ and the correlations measured between the $\hat{y}_{qj}$'s have been very low even for strongly correlated sequences of X_n's (for these experiments the sample sizes were such that $m \geq 200$). Any residual correlations between the group medians could be further reduced by deleting observations between groups, i.e. by discarding observations before starting the next group.

3.3. Theoretical Comparisons

It is interesting to compare the expected confidence interval widths using these three methods under the assumption that each method is valid. Let $m = N/v$ and $m_G = N/vG = m/G$ and assume (without loss of generality) that $p(0;x_p)/m_G f^2(x_p) = 1$. Let $G = 5$ and assume that $\hat{y}_{q1},\ldots,\hat{y}_{qG}$ are i.i.d. normals with mean x_p and variance $p(0;x_p)/m_G f^2(x_p) = 1$. Then the expected ngq confidence interval half-width is 1.163 (Pearson and Hartley (1966), Table 28, p. 190). Consider now the agq method. The t-multiplier for a 93.75% confidence interval with 4 degrees of freedom is 2.562 so, assuming an unbiased estimate of the standard deviation, the expected half-width is $2.562/\sqrt{5} = 1.145$. Finally, assuming the spectral method with $K = 25$ and $d = 2$ produces an unbiased estimate of the point estimates' standard deviation, $(p(0;x_p)/mf^2(x_p))^{1/2} = 1/\sqrt{5}$. Then the confidence intervals expected half-width is $2.214/\sqrt{5} = .99$, since the effective number of degrees of freedom for the variance estimate was shown in Table 1 of Heidelberger and Welch (1981) to be 7.

Thus, given the assumptions, the spectral confidence intervals (with d = 2) will on the average be narrower than both the agq and ngq intervals and the agq intervals will be slightly narrower than the ngq intervals. If a cubic polynomial is used in the spectral method then the spectral confidence intervals can be shown to be somewhat wider than the agq and ngq intervals. These relationships are empirically verified in Section 4.

4. EMPIRICAL STUDIES

All the quantile estimation methods given in the previous section are asymptotically valid under the ϕ-mixing assumption. The crux of the matter though is whether they work with sample sizes required for the usual precisions required in systems simulations, say 10% to 5% as measured by ratio of the confidence interval half-width to the point estimate. In this section we use empirical sampling (simulation) to study the bias and standard deviations of the three quantile point estimators and the confidence interval coverages and widths of the three quantile interval estimators described in Section 3. The tests were conducted on stationary sequences of correlated, exponentially distributed random variables (NEAR(1) and GNEAR(1) processes; see Lawrance and Lewis (1981a) and (1981b)) and on waiting time sequences in heavily congested single server queues. For each process the 0.50, 0.90, 0.99 and 0.999 quantiles were estimated.

4.1. Processes Simulated

The NEAR(1) process $\{X_n\}$ with parameters α and β is defined as follows. Let $\{E_n\}$ be a sequence of i.i.d. exponentials with mean 1 and let $\{I_n\}$ and $\{K_n\}$ be mutually independent sequences of random variables for which $P\{K_n = 1\} = 1 - P\{K_n = (1-\alpha)\beta\} = \delta$, where $\delta = (1-\beta)/(1-(1-\alpha)\beta)$, and $P\{I_n = \beta\} = 1 - P\{I_n = 0\} = \alpha$. Set $X_0 = E_0$; then

$$X_n = I_n X_{n-1} + K_n E_n \quad , \quad n > 0 \tag{4.1}$$

is a Markovian sequence of exponentially distributed random variables with mean one and correlation structure

$$(4.2) \qquad \rho_k = \mathrm{corr}(X_n, X_{n+k}) = (\alpha\beta)^k = \rho^k , \qquad k \geq 0,\ n > 0 .$$

The GNEAR(1) process yields a process with exponential marginals and alternating positive and negative correlations. This process is defined by $X_0 = E_0$ and

$$(4.3) \qquad X_n = K_n E_n + I_n X'_{n-1} , \qquad n > 0 ,$$

where $X'_n = - \ell n(1 - \exp(- X_{n-1}))$. For the GNEAR(1) process the lag one correlation is

$$(4.4) \qquad \rho_1 = (\alpha\beta) r ,$$

where $r = 1 - \pi^2/6 = - 0.6449$ is the maximum negative correlation attainable in a bivariate exponential distribution. Efficient algorithms for generating the NEAR(1) and GNEAR(1) are given in Lawrance and Lewis (1981b).

The second class of processes we considered were (stationary) waiting time sequences in M/G/1 queues, where the service time distribution has a hyperexponential distribution. Specifically, let W_n denote the waiting time of the nth customer and let $\{A_n , n \geq 1\}$ and $\{S_n , n \geq 0\}$ be the i.i.d. sequences of interarrival times and service times respectively. We assume that

(4.5) $$P\{A_n \leq x\} = 1 - \exp(-\lambda x) , \qquad x \geq 0 , \quad \lambda > 0 ,$$

(4.6) $$P\{S_n \leq x\} = \Pi_1\{1 - \exp(-\mu_1 x)\} + \Pi_2\{1 - \exp(-\mu_2 x)\} ,$$

where μ_i and Π_i are greater than or equal to zero, and $\Pi_1 + \Pi_2 = 1$. Let $\mu = (\Pi_1/\mu_1 + \Pi_2/\mu_2)^{-1}$ and let $\rho = \lambda/\mu < 1$ be the traffic intensity. Let W_0 have the stationary waiting time distribution (see Kleinrock (1975), p. 205; this distribution is a probabilistic mixture of an atom at 0 and two exponentials); then the waiting time sequence $\{W_n\}$ is defined by

(4.7) $$W_{n+1} = (W_n + S_n - A_{n+1})^+ \qquad n \geq 0$$

where $x^+ = \max(x,0)$.

We considered five processes in testing the estimators. They are:

(i) NEAR(1) : $\alpha = \beta = 0.95$;

(ii) NEAR(1) : $\alpha = \beta = 0.995$;

(iii) GNEAR(1) : $\alpha = \beta = 0.995$;

(iv) M/M/1 : $\lambda = 9$, $\mu_1 = 10$, $\Pi_1 = 1$, $\rho = 0.90$;

(v) M/G/1 : $\lambda = 9$, $\mu_1 = 2$, $\mu_2 = 18$, $\Pi_1 = 0.10$, $\rho = 0.90$; (the squared coefficient of variation of the service time distribution is 4.556).

The correlation structures of these processes range from moderate positive correlation (NEAR(1), $\alpha = \beta = 0.95$, $\rho_k \approx .9^k$) to extreme positive correlation (NEAR(1), $\alpha = \beta = 0.995$, $\rho_k \approx .99^k$; M/M/1 and M/G/1) to extreme negative correlation (GNEAR(1)). These processes were all simulated using the LLRANDOM II random number generating package (see Lewis and Uribe (1981)).

4.2. Simulation Results

Tables 2-9 report the results of extensive simulation studies of the quantile estimation procedures detailed in earlier Sections. The notation used in the tables is defined as follows. For each process, quantile level p, value v for the max-transform, and run length N, R i.i.d. replications were performed ($R = 200$ for all runs, except runs in which $p = 0.999$, and runs in which $p = .99$ with large values of m). However, the same random number seeds were used for all processes, p, v and N so that different entries in the tables are not independent. Let $\hat{y}_q(r)$, $ngq(x_p, r)$ and $agq(x_p, r)$ denote the realizations on the rth replication of $\hat{y}_q$, $ngq(x_p)$ and $agq(x_p)$ respectively. Let $\bar{y}_q$, $\overline{ngq}$ and $\overline{agq}$ denote the averages over the R replications of $\hat{y}_q(r)$, $ngq(x_p,r)$ and $agq(x_p,r)$ respectively. For example

$$\bar{y}_q = \sum_{r=1}^{R} \hat{y}_q(r)/R \; . \tag{4.8}$$

Let $sd(\bar{y}_q)$, $sd(\overline{ngq})$ and $sd(\overline{agq})$ denote the sample standard deviations of $\bar{y}_q$, $\overline{ngq}$ and $\overline{agq}$ respectively. For example

$$(4.9) \qquad sd^2(\bar{y}_q) = \sum_{r=1}^{R} (\hat{y}_q(r) - \bar{y}_q)^2/R(R-1) \ .$$

These empirical sampling point estimates and their respective estimated standard deviations may be used to study the bias in the quantile estimates derived in this paper. For example, from Table 2, an approximate 90% confidence interval for $E(\hat{y}_q)$ in the NEAR(1) process with $\alpha = \beta = 0.95$, $p = 0.99$, $N = 69000$ and $v = 69$ is $\bar{y}_q \pm 1.645\ sd(\bar{y}_q) = 4.599 \pm 1.645 \times .009$. The true value $x_{.99} = 4.605$ lies within this 90% confidence interval, which is (4.584, 4.614).

(i) Table 2; point estimates of quantiles

Table 2 illustrates the effect of the maximum transformation on $\bar{y}_q$ and its standard deviation, $sd(\bar{y}_q)$. For each process, run length N and $p = 0.90$ and 0.99, $\bar{y}_q$ and $sd(\bar{y}_q)$ were computed using $v = 1$ (no max transform) and v chosen so that $p^v \approx 0.50$. The run lengths N were chosen so that $m = N/V = 1000$ for the NEAR(1) and GNEAR(1) processes and $m = N/v = 2000$ for the M/M/1 and M/G/1 queues. These yield conservative values of m. The only case in which the max-transform introduces apparent bias is the M/M/1 queue with $p = 0.90$, $v = 7$, $N = 14000$; in this case $\bar{y}_p$ differs from x_p by 2.45 times the estimated standard deviation. Since this is the maximum deviation of 20 experiments, it is probably not significant.

The column labeled "Actual sd inflation" is (with p and N fixed) the ratio $sd(\bar{y}_q | p^v \approx 0.50)/sd(\bar{y}_q | v=1)$. This ratio measures the increase in standard deviation in dependent

sequences due to the max-transform. For example, for the NEAR(1) with $\alpha = \beta = 0.995$, $p = 0.99$ and $N = 69000$ this ratio is .0090/.0087 = 1.03 .

The column labeled "Theoretical sd inflation for i.i.d. samples" is the inflation in standard deviation due to the max-transform for a sequence if i.i.d. random variables; see equation (2.8). Recall that the max-transform changes the correlation structure of the process so that its effect on the standard deviation in dependent sequences may differ substantially from its effect in independent sequences. In most cases the actual inflation in variance is very slight and in several cases a small variance reduction is achieved. Only for the M/M/1 queue is the inflation greater than what it would be for an i.i.d. sequence, and even in this case it is, considering the storage and computational savings of the max-transform,an acceptable 1.35 .

From Table 2 we conclude that, for sensibly large values of m , the max-transform introduces very little, if any, bias and that the inflation in variance is modest.

(ii) Tables 3-8; confidence interval estimates

Tables 3-8, in addition to reporting the estimated means and standard deviations of the quantile point estimates, compile the results of confidence interval coverage studies for the three quantile confidence interval procedures. For each process, p , v , N and replication number, confidence intervals for x_p were generated using the spectral method, nested group quantiles, and averaged group quantiles as described in Section 3.

For these studies, G, the number of groups was set at 5. For this value of G, under the assumptions of Section 3 the ngq confidence interval has nominal coverage 0.9375, i.e. $P\{\hat{y}_{q(1)} \leq x_p < \hat{y}_{q(5)}\} = 0.9375$. For comparison purposes we also formed 93.75% confidence intervals using the spectral and agq confidence intervals. Let $CI_r(\hat{y}_q)$, $CI_r(ngq)$ and $CI_r(agq)$ denote the confidence interval on the rth replication using the spectral, ngq and agq methods respectively and let $|CI_r(\hat{y}_q)|$, $|CI_r(ngq)|$ and $|CI_r(agq)|$ denote the widths of these intervals. For each confidence interval procedure Tables 3-8 report the estimated average confidence interval relative half-width (labeled $\overline{hw}/x_p$). For example, for the spectral confidence intervals

$$\overline{hw}/x_p = (1/R) \sum_{i=1}^{R} |CI_r(\hat{y}_q)|/2x_p \,. \tag{4.10}$$

These tables also report the fraction of these confidence intervals which actually contain x_p. This fraction is called an (estimated) 93.75% coverage and it should be close to 0.9375 if valid confidence intervals are being formed. A one sided confidence interval for a coverage, based on the normal approximation to the binomial distribution, can be used to test the hypothesis that the actual coverage is less than 0.9375. For $R = 200$ (Tables 3,4,5,6 and part of 8) estimated coverages less than or equal to 0.91 are significantly low at the 0.90 level while for $R = 100$ (Table 7 and part of 8) estimated coverages less than or equal to 0.89 are significantly low at

the 0.90 level. However, in practice, the importance of coverage in judging the quality of a confidence interval procedure depends very much on the accuracy (as measured by relative half-width) of the confidence intervals. Thus low coverage is not a serious drawback if the procedure generates confidence intervals which are very wide (very inaccurate) and therefore provide very limited information about the quantity being estimated. In fact, the most useful information such a confidence interval often provides is that the run length is too short. As the relative half-width decreases, the importance of coverage for the validity of a procedure increases. We note that it would also be possible to plot the coverage functions for the spectral and agq confidence intervals (Schruben (1980)); however, reporting the individual 93.75% coverages is representative of the coverage function and more compact.

For the spectral method and the M/M/1 and M/G/1 queues the coverages using both degrees $d = 2$ and $d = 3$ are given whereas for the NEAR(1) and GNEAR(1) processes only $d = 2$ is given. For example, from Table 4, the coverages for x_p in the M/M/1 queue with $p = 0.90$, $v = 1$ and $N = 14000$ for the spectral method using degrees $d = 2$ and 3, nested group quantiles and averaged group quantiles are .905, .910, .880 and .870 respectively. The average relative half-widths are .549, .570, .429 and .403 respectively. Thus to increase the precision from its current value of about 50% to a desired value of 10%, we estimate that the sample size would have to be increased by a factor of 25 to a total of 350,000 observations.

It would clearly be impractical to store and sort all of these observations. Thus to increase the precision to an acceptable level requires the max-transform to reduce the sample to a manageable size.

Notice that when the coverages of all methods are valid (for example in Table 5 for all run lengths of the NEAR(1) and GNEAR(1) processes and large run lengths of the queueing processes) that the relative widths of the confidence intervals are generally as predicted in Section 3, namely that

$$\overline{hw}(\text{spectral, d=2}) \leq \overline{hw}(\text{agq}) \leq \overline{hw}(\text{ngq}) \leq \overline{hw}(\text{spectral, d=3}) \ .$$

Tables 3 and 4 list results of experiments for estimating the median and 0.90 quantile respectively without using the max-transform (v = 1). The run lengths used were small enough to accommodate storing and sorting the full sequences.

In Table 3 all coverages are either acceptable (i.e., not significantly less than 0.9375) or nearly acceptable (i.e., wide confidence intervals with coverage close to, but significantly less than 0.9375) with the exception the GNEAR(1) process using the spectral method. The reason for the low coverage in this case is that the sequence of indicator functions $\{I_k(\hat{x}_{.5})\}$ is very nearly a deterministic sequence of alternating zeros and ones due to the strong alternating negative and positive correlation structure of the GNEAR(1) process. This phenomenon appears to be peculiar to the median (the coverage for the 0.90 quantile using the spectral method in Table 4 is acceptable) but should be kept in mind when dealing with processes having this type of correlation structure.

The coverages for the 0.90 quantiles in Table 4 are all acceptable or nearly acceptable. However, $\overline{agq}$ and $\overline{ngq}$ exhibit substantial bias in the M/G/1 queue. In both Tables 3 and 4 notice the very large mean relative half-widths in the M/M/1 and M/G/1 queues, again implying the need for the max-transform to deal with the much larger sample sizes required for estimates of acceptable precision.

Tables 5, 6 and 7 compile results of experiments for estimating the 0.90, 0.99 and 0.999 quantiles, respectively, using the max-transform with v chosen so that $p^v \approx 0.50$.

The ngq and agq confidence interval estimate coverages are generally acceptable throughout these tables. However, as in the case of $v = 1$, $\overline{ngq}$ and $\overline{agq}$ show that there is substantial bias in the ngq and agq point estimates in the M/G/1 queue for small N. For N = 224,000 this bias has disappeared even though the precision (about 20%) is still relatively high. The estimates $\bar{x}_p$ show little, if any, bias. The coverages of the confidence interval estimates in the M/G/1 queue are high despite the bias due to the extreme width of the confidence intervals at the small run lengths.

For the spectral method the coverages for the NEAR(1) and GNEAR(1) processes are acceptable. However, the small sample coverages for the queues, particularly the M/G/1 queue, are disappointing. Even using a cubic polynomial to fit the log of the periodogram does not yield acceptable coverages for the M/G/1 queue when N is small. The low coverage is explained by the fact that the spectrum of the max-transformed variables

$\{Y_k\}$ is still very steep near $f = 0$ (i.e., the Y_i's are highly correlated) even for large values of v. Since $N = mv$, if v is large then m is small and the least squares fits to the log of the periodogram is done over a relatively large frequency range. If $\log(p(f; x_p))$ is very steep then it may not be well approximated by a low order polynomial over a wide frequency range.

Figure 1 illustrates this problem in the case of the NEAR(1) process with $\alpha = \beta = 0.995$. The log of the spectrum of the max-transformed variables ($v = 693$) is seen to be still quite steep near $f = 0$. Compare this to the almost flat log of $p_B(f)$, the spectrum of the process of batch means $\bar{X}_B(k)$, where

$$\bar{X}_B(k) = \frac{1}{B} \sum_{j=(k-1)B+1}^{kB} X_j . \tag{4.11}$$

These batch means would be used to place a confidence interval on $E(X_j)$ (see Heidelberger and Welch (1981)).

The problem can be avoided by making v smaller, and thus m larger. The least squares fit is then done over a narrower frequency range and the coverage improves. This effect is seen by comparing the coverages for $x_{.90}$ in the M/G/1 queue with $N = 14000$ and $v = 1$ (Table 4) to those using $v = 7$ and $N = 14000$ (Table 5). (Again note that the precision is unacceptably high). It is also seen in Table 8 which, for the M/M/1 and M/G/1 queues, compares the coverages for $x_{.99}$

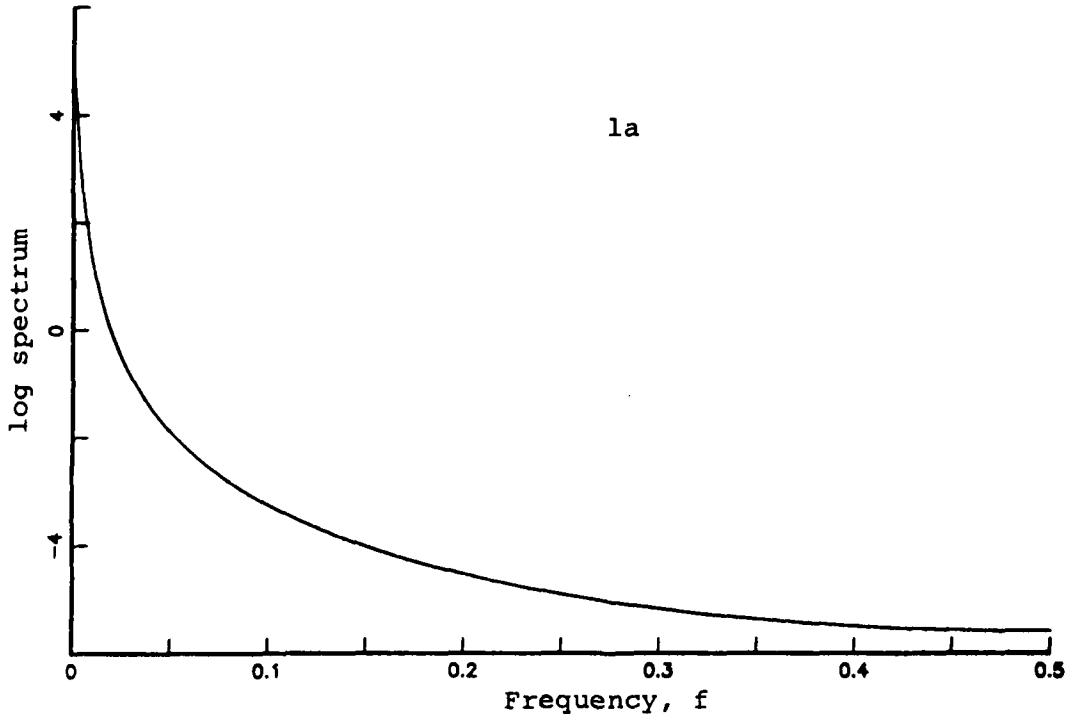


Figure 1. Effect of the maximum transform and batching on the log-spectrum. Figure 1a gives the log of the spectrum of a NEAR(1) process with $\alpha = \beta = 0.995$. Figure 1b shows the log of the spectrum after the maximum transformation with $v = 693$. A small but significant effect can be seen at $f = 0$. Figure 1c shows the log of the spectrum after batching, with batch sizes of $v = 693$. Batching reduces the initial point of the spectrum much more than the maximum transformation.

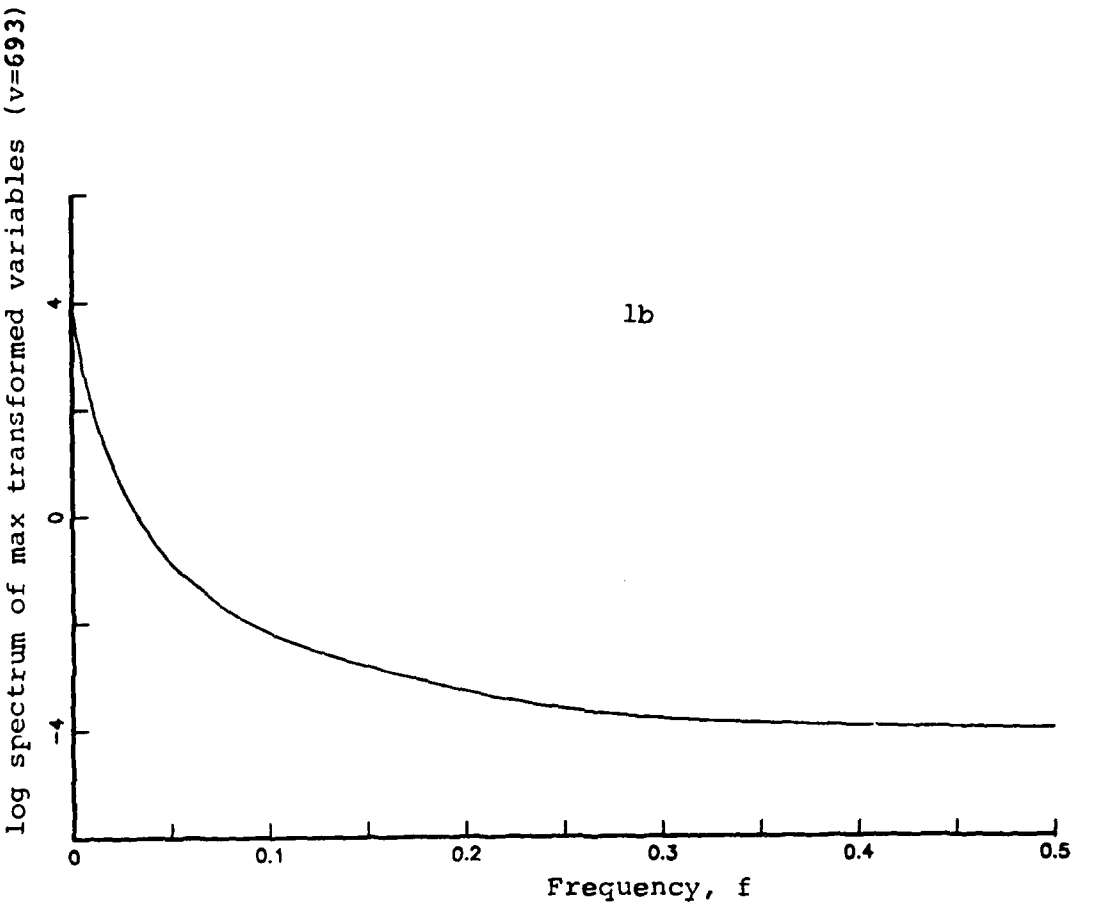


Figure 1. Effect of the maximum transform and batching on the log-spectrum. Figure 1a gives the log of the spectrum of a NEAR(1) process with $\alpha = \beta = 0.995$. Figure 1b shows the log of the spectrum after the maximum transformation with v = 693. A small but significant effect can be seen at f = 0. Figure 1c shows the log of the spectrum after batching, with batch sizes of v = 693. Batching reduces the initial point of the spectrum much more than the maximum transformation.

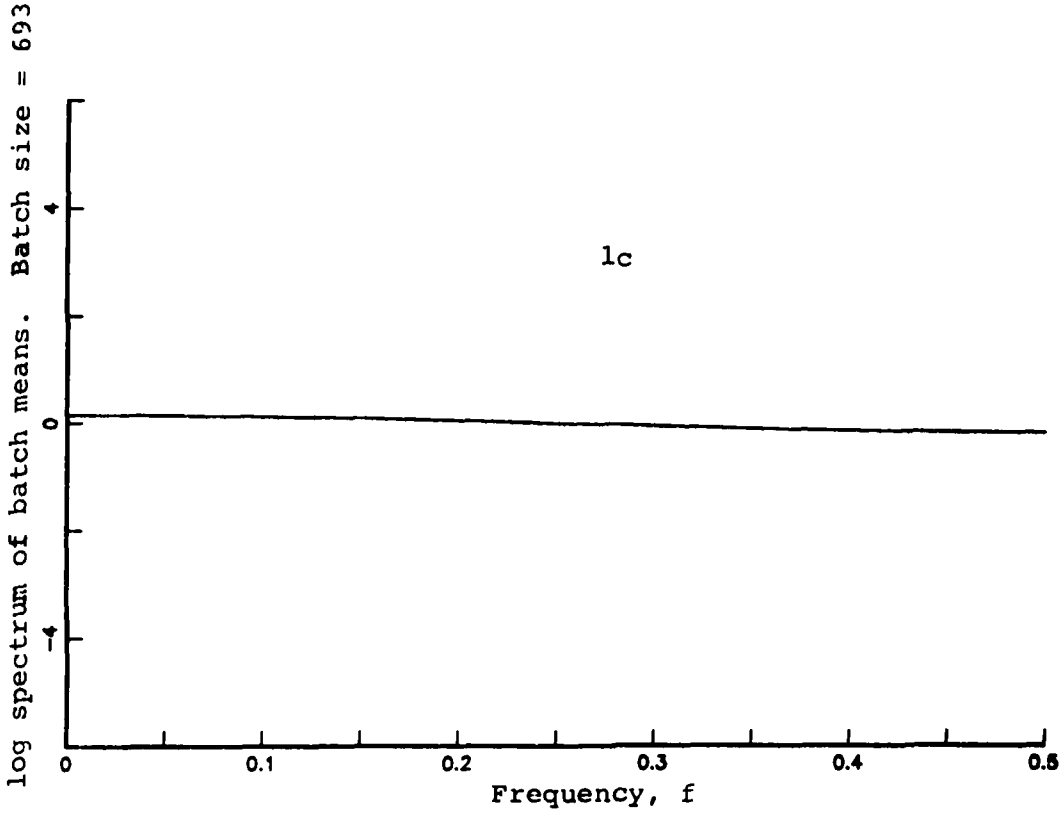


Figure 1. Effect of the maximum transform and batching on the log-spectrum. Figure 1a gives the log of the spectrum of a NEAR(1) process with $\alpha = \beta = 0.995$. Figure 1b shows the log of the spectrum after the maximum transformation with $v = 693$. A small but significant effect can be seen at $f = 0$. Figure 1c shows the log of the spectrum after batching, with batch sizes of $v = 693$. Batching reduces the initial point of the spectrum much more than the maximum transformation.

using $v = 69$ to those using $v = 17$ and $v = 8$, and the coverages for $x_{.999}$ using $v = 693$ to those using $v = 173$ and $v = 86$ (the slight differences in N are to make m highly composite which increases the computational efficiency of the Discrete Fourier Transform used to calculate the periodogram). Thus the full storage and computational savings of the max-transform (i.e. moving the quantile back to the median) cannot be realized in highly correlated sequences without sacrificing confidence interval coverage. *Significant savings can be realized* however. For example the max-transform with $v = 173$ reduces the 1,384,000 observations to a sample size of just 8000 for estimating $x_{.999}$ in the M/G/1 queue.

In Table 8 two simulations are also shown with $N = 5,536,000$ and $v = 173$ for both the M/M/1 and M/G/1 queues for the extreme case of the .999 quantiles. This sample size is large enough to obtain almost 10% precision. To within the precision of the empirical sampling experiment ($R = 100$), it should be noted that all point and confidence interval estimates are valid. Note that the maximum transformation reduces the sample size from 5,536,000 to 32,000!!

5. SUMMARY AND CONCLUSIONS

In this paper we have investigated point and interval estimates for a single quantile in a fixed length sequence of dependent observations. For large sample sizes it becomes impractical to store and sort the entire sequence. For extreme quantiles, these limitations can be overcome by using the maximum transformation which requires storing only a sequence of maxima. This sequence is defined by laying the data out into a v by m array and storing only the maximum element in each column. Storage requirements are thus reduced by a factor of v. Observations at lag m are assumed to be independent and the max-transform changes the problem of estimating the p quantile of the original sequence into one of estimating the p^v quantile of the max-transformed sequence.

Three confidence interval methods which can exploit the sample size reduction produced by the max-transformation were described and tested; the spectral method and two extensions to the method of batch means called nested group quantiles (ngq) and averaged group quantiles (agq). The three methods produced valid confidence intervals if the sample sizes were large enough to produce 10% precision in the estimates.

Problems which were not addressed in this paper and are the subject of ongoing research involve questions of multiple quantile estimation and the incorporation of these fixed run length procedures into sequential procedures for, say, simulation run length control. Among the issues involved here are:

(i) Testing the adequacy of the spacing parameter m so that X_k and X_{k+m} are approximately independent.

(ii) Organizing the data so that m can be increased as the sample size increases.

(iii) Organizing the data so that multiple quantiles can be efficiently estimated, for example, estimating a 0.50 and 0.999 quantile for the same set of observations (this probably involves the use of a max-min transformation).

(iv) Determining the sensitivity of these schemes to an initial transient which is often encountered in simulation.

Acknowledgements

The work of P. A. W. Lewis was supported by the Office of Naval Research under Grant NR-42-284 and by the IBM Research Laboratories under a consulting agreement.

Table 2. Comparison of point estimates and standard deviations of point estimates using full sample order statistic and max-transformed sample order statistic (R = 200 replications). Here $q = p^V$. The fact that the actual sd inflation is generally less than the theoretical sd inflation for i.i.d. sequences is because of the decrease in correlation in the series induced by the maximum transformation.

Process	p	x_p	N	V	$\bar{y}_q$	$sd(\hat{y}_q)$	Actual sd inflation	Theoretical sd inflation for i.i.d. sequences
NEAR(1)	0.90	2.303	7,000	1	2.304	.0104		
$\alpha=\beta=0.95$	0.90	2.303	7,000	7	2.320	.0106	1.02	1.18
	0.99	4.605	69,000	1	4.610	.0087		
	0.99	4.605	69,000	69	4.599	.0090	1.03	1.20
NEAR(1)	0.90	2.303	7,000	1	2.344	.0325		
$\alpha=\beta=0.995$	0.90	2.303	7,000	7	2.340	.0356	1.10	1.18
	0.99	4.605	69,000	1	4.540	.0275		
	0.99	4.605	69,000	69	4.607	.0299	1.09	1.20
GNEAR(1)	0.90	2.303	7,000	1	2.282	.0132		
$\alpha=\beta=0.995$	0.90	2.303	7,000	7	2.291	.0121	0.92	1.18
negative correlation	0.99	4.605	69,000	1	4.605	.0046		
	0.99	4.605	69,000	69	4.600	.0050	1.09	1.20
M/M/1	0.90	2.197	14,000	1	2.222	.0277		
$\rho=0.90$	0.90	2.197	14,000	7	2.289	.0375	1.35	1.18
	0.99	4.500	138,000	1	4.514	.0368		
	0.99	4.500	138,000	69	4.558	.0455	1.24	1.20
M/G/1	0.90	6.311	14,000	1	6.466	.1682		
$\rho=0.90$	0.90	6.311	14,000	7	6.572	.1759	1.05	1.18
	0.99	13.130	138,000	1	13.182	.2077		
	0.99	13.130	138,000	69	13.398	.1975	0.95	1.20

Table 3. Point estimates, standard deviations of point estimates, 93.75% coverages, and average confidence interval half-width/x_p for the three methods. Here p = 0.50, v = 1, G = 5, and R, the number of replications, is 200.

Process	x_p	N	Spectral Method			Nested Group Quantiles		Averaged Group Quantiles	
				degree of polynomial 3	degree of polynomial 2				
			$\bar{x}_p$ sd($\bar{x}_p$)	coverage $\overline{hw}/x_p$	coverage $\overline{hw}/x_p$	$\overline{ngq}$ sd($\overline{ngq}$)	coverage $\overline{hw}/x_p$	$\overline{ngq}$ sd($\overline{agq}$)	coverage $\overline{hw}/x_p$
NEAR(1) $\alpha=\beta=0.95$	0.693	7,000	.696 (.003)		.935 (.149)	.700 (.004)	.935 (.180)	.706 (.003)	.945 (.167)
NEAR(1) $\alpha=\beta=0.995$	0.693	7,000	.735 (.012)		.865 (.517)	.744 (.015)	.905 (.598)	.809 (.013)	.910 (.565)
GNEAR(1) $\alpha=\beta=0.995$ (negative correlation)	0.693	7,000	.692 (.001)		.795 (.006)	.693 (.002)	.945 (.112)	.689 (.002)	.985 (.102)
M/M/1 $\rho=0.90$	0.588	14,000	.601 (.006)	.950 (.452)	.950 (.352)	.598 (.007)	.945 (.443)	.640 (.008)	.940 (.413)
M/G/1 $\rho=0.90$	1.545	14,000	1.607 (.033)	.950 (.878)	.940 (.647)	1.562 (.036)	.930 (1.013)	1.943 (.055)	.940 (.961)

Table 4. Point estimates, standard deviations of point estimates, 93.75% coverages, and average confidence interval half-width/x_p for the three methods, no maximum transform. Here p = 0.90, v = 1, G = 5 and R, the number of replications, is 200.

Process	x_p	N	Spectral Method			Nested Group Quantiles		Averaged Group Quantiles	
				degree of polynomial 3	2				
			$\bar{x}_p$	coverage	coverage	$\overline{ngq}$	coverage	$\overline{ngq}$	coverage
			$sd(\bar{x}_p)$	$\overline{hw}/x_p$	$\overline{hw}/x_p$	$sd(\overline{ngq})$	$\overline{hw}/x_p$	$sd(\overline{agq})$	$\overline{hw}/x_p$
NEAR(1) $\alpha=\beta=0.95$	2.303	7,000	2.304 (.010)		.930 (.129)	2.287 (.013)	.925 (.156)	2.296 (.010)	.915 (.145)
NEAR(1) $\alpha=\beta=0.995$	2.303	7,000	2.344 (.032)		.935 (.468)	2.189 (.038)	.955 (.469)	2.264 (.030)	.940 (.435)
GNEAR(1) $\alpha=\beta=0.995$ (negative correlation)	2.303	7,000	2.305 (.010)		.925 (.141)	2.282 (.013)	.955 (.181)	2.267 (.011)	.955 (.168)
M/M/1 $\rho=0.90$	2.197	14,000	2.222 (.028)	.910 (.570)	.905 (.549)	2.066 (.027)	.880 (.429)	2.229 (.027)	.870 (.403)
M/G/1 $\rho=0.90$	6.311	14,000	6.466 (.168)	.890 (1.069)	.845 (.669)	5.266 (.090)	.865 (.496)	5.843 (.099)	.815 (.466)

Table 5. Point estimates, standard deviations of point estimates, 93.75% coverages, and average confidence interval half-width/x_p for the three methods. As in Table 4, p = .90, but here v = 7; thus $p^v \approx 0.50$. Also G = 5 and R, the number of replications, is 200.

Process	x_p	N	Spectral Method			Nested Group Quantiles		Averaged Group Quantiles	
				degree of polynomial 3	2				
			$\bar{y}_q$	coverage	coverage	$\overline{ngq}$	coverage	$\overline{ngq}$	coverage
			$sd(\bar{y}_q)$	$\overline{hw}/x_p$	$\overline{hw}/x_p$	$sd(\overline{ngq})$	$\overline{hw}/x_p$	$sd(\overline{agq})$	$\overline{hw}/x_p$
NEAR(1) $\alpha=\beta=0.95$	2.303	7,000	2.320 (.011)		.945 (.156)	2.310 (.015)	.940 (.179)	2.328 (.012)	.940 (.164)
		28,000	2.304 (.006)		.940 (.075)	2.303 (.007)	.920 (.089)	2.305 (.006)	.915 (.083)
		112,000	2.300 (.003)		.930 (.037)	2.296 (.003)	.920 (.044)	2.300 (.003)	.935 (.040)
NEAR(1) $\alpha=\beta=0.995$	2.303	7,000	2.340 (.036)		.800 (.342)	2.253 (.038)	.940 (.519)	2.414 (.033)	.945 (.488)
		28,000	2.311 (.017)		.935 (.246)	2.297 (.020)	.960 (.279)	2.350 (.017)	.940 (.260)
		112,000	2.289 (.008)		.970 (.127)	2.301 (.010)	.945 (.142)	2.309 (.008)	.945 (.131)
GNEAR(1) $\alpha=\beta=0.995$ (negative correlation)	2.303	7,000	2.291 (.012)		.960 (.199)	2.316 (.013)	.935 (.192)	2.287 (.012)	.940 (.178)
		28,000	2.306 (.006)		.940 (.078)	2.312 (.007)	.940 (.097)	2.299 (.006)	.935 (.090)
		112,000	2.304 (.003)		.925 (.040)	2.306 (.003)	.925 (.047)	2.304 (.003)	.935 (.043)
M/M/1 $\rho=0.90$	2.197	14,000	2.289 (.037)	.940 (.500)	.815 (.306)	2.206 (.033)	.955 (.451)	2.331 (.032)	.930 (.422)
		56,000	2.231 (.016)	.975 (.325)	.950 (.240)	2.220 (.020)	.945 (.296)	2.316 (.020)	.935 (.274)
		224,000	2.206 (.008)	.965 (.142)	.950 (.116)	2.204 (.009)	.955 (.138)	2.227 (.008)	.945 (.127)
M/G/1 $\rho=0.90$	6.311	14,000	6.572 (.176)	.695 (.399)	.525 (.219)	5.657 (.123)	.865 (.500)	6.112 (.110)	.860 (.472)
		56,000	6.373 (.079)	.950 (.471)	.870 (.300)	6.184 (.079)	.940 (.444)	6.680 (.085)	.935 (.417)
		224,000	6.388 (.042)	.970 (.260)	.955 (.209)	6.367 (.047)	.965 (.243)	6.559 (.047)	.960 (.227)

Table 6. Point estimates, standard deviations of point estimates, 93.75% coverages, and average confidence interval half-width/x_p for the three methods. Here p = 0.99, v = 69, so that $q = p^v \approx 0.50$. Also G = 5 and the number of replications, R, is 200. However R = 100 if the run is marked †.

Process	x_p	N	Spectral Method			Nested Group Quantiles		Averaged Group Quantiles	
				degree of polynomial 3	2				
			$\bar{y}_q$	coverage	coverage	$\overline{ngq}$	coverage	$\overline{ngq}$	coverage
			$sd(\bar{y}_q)$	$\overline{hw}/x_p$	$\overline{hw}/x_p$	$sd(\overline{ngq})$	$\overline{hw}/x_p$	$sd(\overline{agq})$	$\overline{hw}/x_p$
NEAR(1) $\alpha=\beta=0.95$	4.605	69,000	4.599 (.009)		.960 (.063)	4.606 (.011)	.920 (.073)	4.627 (.009)	.915 (.067)
		138,000	4.608 (.006)		.945 (.043)	4.604 (.007)	.950 (.054)	4.608 (.006)	.950 (.050)
		276,000	4.604 (.004)		.930 (.031)	4.603 (.005)	.945 (.037)	4.608 (.005)	.950 (.034)
NEAR(1)	4.605	69,000	4.607 (.030)		.890 (.180)	4.606 (.032)	.965 (.248)	4.692 (.029)	.965 (.229)
		138,000	4.653 (.023)		.930 (.140)	4.653 (.027)	.905 (.173)	4.685 (.024)	.925 (.160)
		276,000	4.620 (.014)		.940 (.099)	4.613 (.016)	.945 (.111)	4.629 (.014)	.935 (.102)
GNEAR(1) $\alpha=\beta=0.995$	4.605	69,000	4.600 (.005)		.955 (.035)	4.596 (.006)	.950 (.040)	4.594 (.005)	.940 (.037)
		138,000	4.605 (.004)		.920 (.023)	4.605 (.005)	.925 (.028)	4.601 (.004)	.910 (.026)
(negative correlation)		276,000	4.608 (.003)		.950 (.017)	4.604 (.003)	.930 (.020)	4.605 (.003)	.935 (.019)
M/M/1	4.500	138,000	4.558 (.045)	.865 (.262)	.730 (.153)	4.351 (.035)	.920 (.267)	4.538 (.035)	.905 (.250)
		276,000	4.513 (.027)	.955 (.236)	.895 (.148)	4.472 (.032)	.930 (.222)	4.598 (.028)	.930 (.207)
		552,000	4.525 (.020)	.955 (.190)	.945 (.132)	4.511 (.024)	.945 (.178)	4.608 (.023)	.935 (.166)
†		2,208,000	4.481 (.012)	.980 (.084)	.960 (.067)	4.495 (.015)	.960 (.076)	4.510 (.013)	.950 (.070)
M/G/1	13.130	138,000	13.398 (.198)	.675 (.243)	.490 (.129)	11.713 (.125)	.820 (.288)	12.174 (.119)	.815 (.270)
		276,000	13.359 (.154)	.845 (.270)	.635 (.155)	12.377 (.107)	.920 (.287)	12.934 (.106)	.900 (.267)
		552,000	13.271 (.106)	.935 (.276)	.815 (.164)	12.927 (.103)	.910 (.265)	13.480 (.101)	.915 (.247)
†		2,208,000	13.102 (.062)	.960 (.163)	.930 (.123)	13.010 (.075)	.920 (.133)	13.241 (.066)	.900 (.125)

Table 7. Point estimates, standard deviations of point estimates, 93.75% coverages, and average confidence interval half-width/x_p for the three methods. Here p = 0.999, v = 693, so that $q = p^v \simeq 0.50$. Also G = 5 and the number of replications, R, is 100.

Process	x_p	N	Spectral Method: $\bar{y}_q$ / $sd(\bar{y}_q)$	Spectral Method: degree of polynomial 3: coverage / $\overline{hw}/x_p$	Spectral Method: degree of polynomial 2: coverage / $\overline{hw}/x_p$	Nested Group Quantiles: $\overline{ngq}$ / $sd(\overline{ngq})$	Nested Group Quantiles: coverage / $\overline{hw}/x_p$	Averaged Group Quantiles: $\overline{ngq}$ / $sd(\overline{agq})$	Averaged Group Quantiles: coverage / $\overline{hw}/x_p$
NEAR(1) $\alpha=\beta=0.95$	6.908	693,000	6.923 (.011)		.940 (.034)	6.908 (.012)	.940 (.038)	6.912 (.010)	.940 (.035)
NEAR(1) $\alpha=\beta=0.995$	6.908	693,000	6.950 (.033)		.960 (.117)	6.937 (.040)	.920 (.138)	7.018 (.037)	.920 (.128)
GNEAR(1) $\alpha=\beta=0.995$ (negative correlation)	6.908	693,000	6.906 (.005)		.930 (.015)	6.899 (.006)	.930 (.018)	6.899 (.005)	.930 (.016)
M/M/1 $\rho=0.90$	6.800	1,386,000	6.835 (.058)	.900 (.181)	.730 (.108)	6.648 (.046)	.930 (.172)	6.790 (.046)	.940 (.160)
M/G/1 $\rho=0.90$	19.94	1,386,000	20.693 (.324)	.700 (.161)	.520 (.093)	18.589 (.171)	.890 (.211)	19.063 (.157)	.840 (.198)

Table 8. M/M/1 queue. Point estimates, standard deviations of point estimates, 93.75% coverages, and average confidence interval half-widths/x_p comparing $q = p^v \approx 0.50$, $q = p^v \approx 0.84$ and $q = p^v \approx 0.92$ for different sample sizes. Here $p = 0.99$ and $p = 0.999$ and $G = 5$. The number of replications is $R = 200$ for runs marked * , and $R = 100$ for runs marked † .

Process		p x_p	N	Spectral Method $\bar{y}_q$ $sd(\bar{y}_q)$	degree of polynomial 3 coverage $\overline{hw}/x_p$	degree of polynomial 2 coverage $\overline{hw}/x_p$	Nested Group Quantiles $\overline{ngq}$ $sd(\overline{ngq})$	coverage $\overline{hw}/x_p$	Averaged Group Quantiles $\overline{ngq}$ $sd(\overline{agq})$	coverage $\overline{hw}/x_p$
M/M/1	*	0.99	138,000	4.558	.865	.730	4.351	.920	4.538	.905
ρ=0.90		4.500	69	(.045)	(.262)	(.153)	(.035)	(.267)	(.035)	(.250)
	*	0.99	136,000	4.546	.925	.880	4.258	.895	4.409	.890
		4.500	17	(.040)	(.375)	(.246)	(.029)	(.240)	(.029)	(.222)
	*	0.99	128,000	4.596	.900	.870	4.317	.895	4.470	.875
		4.500	8	(.042)	(.384)	(.306)	(.035)	(.260)	(.037)	(.243)
	†	0.999	1,386,000	6.835	.900	.730	6.648	.930	6.790	.940
		6.800	693	(.058)	(.181)	(.108)	(.046)	(.172)	(.046)	(.160)
	†	0.999	1,384,000	6.855	.910	.870	6.578	.930	6.732	.910
		6.800	173	(.059)	(.284)	(.178)	(.055)	(.178)	(.050)	(.165)
	†	0.999	1,376,000	6.856	.940	.890	6.657	.900	6.754	.880
		6.800	86	(.056)	(.232)	(.181)	(.054)	(.156)	(.049)	(.144)
	†	0.999	5,536,000	6.797	.970	.950	6.743	.930	6.833	.940
		6.800	173	(.024)	(.112)	(.088)	(.027)	(.108)	(.028)	(.100)
M/G/1	*	0.99	138,000	13.398	.675	.490	11.713	.820	12.174	.815
ρ=0.90		13.130	69	(.198)	(.243)	(.129)	(.125)	(.288)	(.119)	(.270)
	*	0.99	136,000	13.240	.850	.725	11.467	.795	12.011	.755
		13.130	17	(.202)	(.462)	(.274)	(.125)	(.294)	(.124)	(.277)
	*	0.99	128,000	13.011	.805	.730	11.336	.755	11.762	.720
		13.130	8	(.200)	(.758)	(.489)	(.126)	(.286)	(.115)	(.266)
	†	0.999	1,386,000	20.693	.700	.520	18.589	.890	19.063	.840
		19.94	693	(.324)	(.161)	(.093)	(.171)	(.211)	(.157)	(.198)
	†	0.999	1,384,000	19.951	.890	.800	18.188	.880	18.638	.880
		19.94	173	(.230)	(.294)	(.170)	(.159)	(.198)	(.137)	(.185)
	†	0.999	1,376,000	19.844	.820	.780	18.053	.800	18.621	.770
		19.94	86	(.239)	(.385)	(.238)	(.176)	(.198)	(.159)	(.183)
	†	0.999	5,536,000	20.152	.950	.920	19.511	.920	19.979	.950
		19.94	173	(.146)	(.211)	(.156)	(.138)	(.151)	(.128)	(.142)

REFERENCES

Babu, G. J. and Singh, K. (1978). "On deviations between empirical and quantile processes for mixing random variables." J. of Multivariate Analysis 8, 532-549.

Billingsley, P. (1968). "Convergence of Probability Measures." Wiley, New York.

Blomqvist, N. (1967). "The covariance function of the M/G/1 queueing system," Skand. Akt. Tidskr., 50, 157-174.

Connover, W. J. (1980). "Practical Nonparametric Statistics, 2n ed." Wiley, New York.

David, H. (1970). "Order Statistics." Wiley, New York.

Fishman, G. S. (1978). "Principles of Discrete Event Simulation." Wiley-Interscience, New York.

Goodman, A. S., Lewis, P. A. W., and Robbins, H. E. (1971). "Simultaneous estimation of large numbers of extreme quantiles in simulation experiments," Naval Postgraduate School Technical Report NPS-55Lw71122A.

Heidelberger, P. and Welch, P. D. (1980). "On the statistical control of simulation run length." IBM Research Report RC8571, Yorktown Heights, NY, to appear in IBM J. Res. Develop. (1981).

Heidelberger, P. and Welch, P. D. (1981). "A spectral method for confidence interval generation and run length control in simulations." Comm ACM 24, 233-245.

Iglehart, D. L. (1976). "Simulating stable stochastic systems VI: quantile estimation," J. Assoc. Comput. Mach. 23, 347-60.

Kleinrock, L. (1975). "Queueing Systems, Volume 1: Theory." John Wiley and Sons, New York.

Law, A. M. and Carson, J. S. (1979). "A Sequential Procedure for Determining the Length of a Steady-State Simulation," Opns. Res. 27, 5, p. 1011-1025.

Lawrance, A. J. and Lewis, P. A. W. (1981a). "A new autoregressive time series model in exponential variables (NEAR(1))." J. Appl. Prob. (to appear).

Lawrance, A. J. and Lewis, P. A. W. (1981b). "Generation of some first-order autoregressive Markovian sequences of positive random variables with given marginal distributions." To appear in Proc. App. Prob./Comp. Science Conf., R. Disney (ed.).

Lewis, P. A. W. and Uribe, L. (1981). "The new Naval Postgraduate School Random Number Package-LLRANDOM II," Naval Postgraduate School Report NPS 55-81-005, Monterey, CA.

Moore, L. R. (1979). "Quantile Estimation in Regenerative Processes." Technical Report 79-10. Curriculum in Operations Research and Systems Analysis. University of North Carolina, Chapel Hill, NC.

Parzen, E. (1962). "On estimation of a probability density function and mode," Annals of Mathematical Statistics, 33, 1065-1076.

Pearson, E. S. and Hartley, H. O. (1966). "Biometrika Tables For Statisticians." Cambridge University Press, England.

Robbins, H. E. and Monro, S. (1951). "A stochastic approximation method." Annals of Mathematical Statistics, 22, 400-407.

Robinson, D. W. (1975). "Non-parametric quantile estimation through stochastic approximation." Ph.D. Thesis, Naval Postgraduate School, Monterey, CA.

Rosenblatt, M. (1956). "Remarks on some nonparametric estimates of a density function," Annals of Mathematical Statistics, 27, 832-835.

Schruben, L. W. (1980). "A coverage function for interval estimators of simulation response." Management Sci. 26, 18-27.

Seila, A. F. (1976). "Quantile estimation methods in discrete event simulations of regenerative systems." Technical Report No. 76-12, Curr. in Operations Research, Univ. N. Carolina, Chapel Hill.

Sen, P. K. (1972). "On the Bahadur representation of sample quantiles for sequences of ϕ-mixing random variables." J. of Multivariate Analysis 2, 77-95.

DISTRIBUTION LIST

	NO. OF COPIES
Defense Technical Information Center Cameron Station Alexandria, VA 22314	2
Library, Code 0142 Naval Postgraduate School Monterey, CA 93940	2
Dean of Research Code 012 Naval Postgraduate School Monterey, CA 93940	1
Library, Code 55 Naval Postgraduate School Monterey, CA 93940	1
Professor P. A. W. Lewis Code 55Lw Naval Postgraduate School Monterey, CA 93940	220

END

DATE
FILMED

11-81

DTIC