

IDENTIFIABILITY IN GI/G/k QUEUES

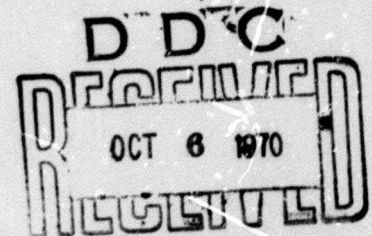
by
SHELDON M. ROSS

AD712066

OPERATIONS
RESEARCH
CENTER

COLLEGE OF ENGINEERING
UNIVERSITY OF CALIFORNIA • BERKELEY

Reproduced by the
CLEARINGHOUSE
for Federal Scientific & Technical
Information Springfield Va. 22151

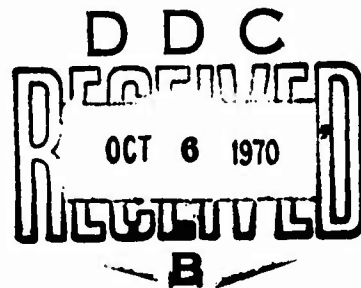


This document has been approved
for public release and sale; its
distribution is unlimited.

IDENTIFIABILITY IN GI/G/k QUEUES

by

Sheldon M. Ross
Department of Industrial Engineering
and Operations Research
University of California, Berkeley



AUGUST 1970

This document has been approved
for public release and sale; its
distribution is unlimited.

ORC 70-26

This research has been supported by the U.S. Army Research Office-Durham under Contract DA-31-124-ARO-D-331 with the University of California. Reproduction in whole or in part is permitted for any purpose of the United States Government.

ABSTRACT

Consider a queueing system in which customers arrive in accordance with a renewal process having an unknown distribution F , and in which the service times are independent and have unknown distribution G . We assume that there are $k(k \leq \infty)$ servers.

Let $C(t)$ denote the number of customers in the system at time t . It is shown that F and G are identifiable from $\{C(t), t \geq 0\}$ if either G or F is continuous, if $F(x) < 1$ for all x , and if the number of busy periods is infinite almost surely. Secondly, it is shown that F and G are identifiable if G is not lattice and the queue size a.s. converges to infinity. The second result is obtained by first considering the related problem: if $N_i(t)$, $i = 1, \dots, k$, are independent renewal processes with the same distribution G , then construct an a.s. convergent estimate of G based only on $\{N(t), t \geq 0\}$, where

$$N(t) = \sum_{i=1}^k N_i(t).$$

IDENTIFIABILITY IN $G1/G/k$ QUEUES

by

Sheldon M. Ross

0. INTRODUCTION AND SUMMARY

Consider a queuing system in which customers arrive in accordance with a renewal process having an interarrival distribution F , and in which the service times of customers are independent and have distribution G . Moreover, suppose that there are k ($k \leq \infty$) servers and that an arriving customer is immediately served if he finds one of the servers free, and if not then he joins the queue.

When a customer completes service, he is immediately replaced by a customer waiting in the queue. However, we do not assume that the queue discipline (whether LIFO, FIFO, or anything else) is known.

Let $C(t)$ denote the number of customers in the system at time t . The problem of interest is to determine whether F and G are identifiable from a single sample path of the process $\{C(t), t \geq 0\}$. In Section 1, we show that F and G are identifiable if either G or F is continuous, if $F(x) < 1$ for all x , and if the number of busy periods is a.s. infinite. In Section 2, we show that F and G are identifiable if G is not lattice and if the queue size a.s. converges to infinity. This is done by first considering the following problem which is of independent interest: If $N_i(t)$, $i = 1, \dots, k$, are independent renewal processes with the same distribution G , then construct an a.s. convergent estimate of G based only on the process $\{N(t), t \geq 0\}$, where $N(t) = \sum_{i=1}^k N_i(t)$.

The problem of identifying F and G has received some attention in the literature. In [1], Brown considered the problem of estimating G from a single sample path of $\{C(t), t \geq 0\}$ for the case $M/G/\infty$ (that is, where F is assumed exponential and $k = \infty$). By a clever use of the ergodic theorem, he showed that his estimate a.s. converged (pointwise) uniformly to G , which, of course, shows that G is identifiable. Brown also raised the question of the identifiability of

G from $\{C(t), t \geq 0\}$ for the $GI/G/\infty$ system. In [3], Kendall and Lewis show that in a $GI/G/\infty$ queue if one is given the output epochs along with the serial number of the input corresponding to each output then the input distribution is completely identifiable and the output distribution is identifiable up to a location parameter.

1. RESULTS WHEN BUSY PERIODS ARE A.S. FINITE

We say that a busy period begins when a customer arrives and finds the system empty. In this section we shall suppose that a busy period is a.s. of finite length. From results of Kiefer-Wolfowitz [4,5] it follows that this will be the case if $\mu_G < k\mu_F$ and $F(x) < 1$ for all x , where μ_G and μ_F are the mean service and interarrival times.

Call a customer a B-customer if there are no other customers in the system when he arrives. Let

$$I_n = \begin{cases} 1 & \text{if } n^{\text{th}} \text{ B-customer completes his service before} \\ & \text{a new customer arrives} \\ 0 & \text{otherwise .} \end{cases}$$

Then, by the strong law of large numbers it follows that

$$(1) \quad \frac{I_1 + \dots + I_n}{n} \xrightarrow{\text{a.s.}} \int_0^{\infty} G(y) dF(y) .$$

Now from the set of all B-customers let us consider only those for which there is an arrival within a time x following their arrival. Call these B_x -customers and let

$$J_n = \begin{cases} 1 & \text{if } n^{\text{th}} B_x\text{-customer completes his service before} \\ & \text{a new customer arrives} \\ 0 & \text{otherwise .} \end{cases}$$

Then from the strong law of large numbers it follows that either $F(x) = 0$ and there are no B_x -customers or else

$$(2) \quad \frac{J_1 + \dots + J_n}{n} \xrightarrow{a.s.} \frac{\int_0^x G(y) dF(y)}{F(x)}.$$

Now, from the set of B-customers consider those who finish before a new customer arrives, and let their service times be $X_1, X_2, \dots$. If we let $H_n(x)$ be the empirical c.d.f. of $X_1, \dots, X_n$, then

$$(3) \quad H_n(x) \xrightarrow{a.s.} H(x) = \frac{\int_0^x G(y) dF(y) + G(x)(1-F(x))}{\int_0^\infty G(y) dF(y)}.$$

Finally let Y_i , $i \geq 1$, be the successive interarrival times and let F_n be the empirical c.d.f. of $Y_1, \dots, Y_n$. Then

$$(4) \quad F_n(x) \xrightarrow{a.s.} F(x).$$

Theorem 1:

If

- (i) the number of busy periods is a.s. infinite
- (ii) either F or G is continuous
- (iii) $F(x) < 1$ for all x

then F and G are identifiable from $\{C(t), t \geq 0\}$.

Proof:

From the continuity assumption (ii) it follows that an arrival and departure may not occur simultaneously, and as a result, the left sides of (1) - (4) are all obtainable from the data $\{C(t), t \geq 0\}$. Hence, if

$F(x) \neq 0$ then the right sides of (1) - (4) are identifiable and the result follows

since $G(x) = \frac{H(x) \int_0^{\infty} G(y) dF(y) - \int_0^x G(y) dF(y)}{1-F(x)}$. On the other hand, if $F(x) = 0$ then the result follows from (1), (3) and (4).

Q.E.D.

Remark:

The purpose of the continuity assumption (ii) is that it enables us to determine the arrival and departure epochs from the data $\{G(t), t \geq 0\}$. If the experimenters data consists of the arrival and departure epochs then this condition is not necessary.

2. THE CASE $k\mu_G > \mu_F$

In this section we suppose that G is nonlattice, either G or F is continuous and $\mu_G > k\mu_F$. Before showing that F and G are identifiable in this case we first consider the following related problem.

Related Problem:

Consider k independent stationary renewal processes $\{N_i(t)\}, i = 1, \dots, k$, for which the time until the first renewal has a distribution $G_e(x) = \int_0^x \frac{(1-G(x))dx}{\mu_G}$, and the remaining interarrival times have distribution G , where G is assumed to be nonlattice. Let $N(t) = \sum_{i=1}^k N_i(t)$. The problem is to estimate G from the data $\{N(t), t \geq 0\}$.

Let E_n denote the time from n until the next jump of the $N(t)$ process, and let

$$I_n = \begin{cases} 1 & E_n > x \\ 0 & E_n \leq x \end{cases}.$$

Lemma 1:

$$1/n \sum_{j=0}^{n-1} I_j \xrightarrow{a.s.} \left(\int_x^\infty \frac{-G(x)}{\mu_G} dx \right)^k.$$

Proof:

From known renewal theory results it follows that $\{E_n, n \geq 0\}$ is a stationary process. Also, by known renewal results, it follows that the distribution of $\{E_n, n \geq m\}$ given the values $E_n, n \leq j$, converges as $m \rightarrow \infty$ to the distribution of $\{E_n, n \geq 0\}$ regardless of the values $E_n, n \leq j$. This implies that the tail σ -field of $\{E_n, n \geq 0\}$ is independent of $\{E_n, n \leq j\}$ and

hence every tail event has probability 0 or 1 which implies that the E_n process is ergodic. Since the E_n process is stationary and ergodic, so is the I_n process. The result then follows from the ergodic theorem since

$$E I_0 = \left(\int_x^\infty \frac{1-G(x)}{\mu_G} dx \right)^k . \quad \text{Q.E.D.}$$

Now, define T_n to be the time between the $(n-1)$ st and the n^{th} jump of the $\{N(t), t \geq 0\}$ process, and let

$$J_n = \begin{cases} 1 & T_n > x \\ 0 & T_n \leq x . \end{cases}$$

Lemma 2:

$$1/n \sum_{i=0}^{n-1} J_i \xrightarrow{a.s.} (1 - G(x)) \left(\int_x^\infty \frac{1-G(x)}{\mu_G} dx \right)^{k-1} .$$

Proof:

Let $S_{n,j}$ denote the time of the n^{th} event of the process $N_j(t)$, $j = 1, \dots, k$, $n > 0$. Let $T_{n,j}$ be the time from $S_{n,j}$ to the next jump in the $N(t)$ process. Now because of the independence and stationarity of the $N_j(t)$ processes, $1 \leq j \leq k$, it follows exactly as in Lemma 1 that for each j , $1 \leq j \leq k$, $\{T_{n,j}, n \geq 1\}$ is a stationary ergodic process. Letting

$$J_{nj} = \begin{cases} 1 & T_{nj} > x \\ 0 & T_{nj} \leq x \end{cases} \quad \text{we have by the ergodic theorem that}$$

$$1/n \sum_{i=0}^{n-1} J_{ij} \xrightarrow{a.s.} (1 - G(x)) \left(\int_x^\infty \frac{1-G(x)}{\mu_G} dx \right)^{k-1}, \quad j = 1, 2, \dots, k .$$

The result now follows since $\{J_n, n \geq 1\}$ is composed of the k subsequences $\{J_{nj}, n \geq 1\}, 1 \leq j \leq k$. Q.E.D.

Remark:

In the proof of Lemma 2 it is necessary to introduce the $T_{n,j}$ since the process $\{T_n, n \geq 1\}$ is not a stationary process.

From Lemmas 1 and 2 we obtain

Theorem 2:

$$\frac{1/n \sum_{j=0}^{n-1} J_1}{\left(1/n \sum_{j=0}^{n-1} I_j\right)^{\frac{k-1}{k}}} a.s. \rightarrow 1 - G(x) \text{ as } n \rightarrow \infty,$$

and hence $G(x)$ is identifiable from $\{N(t), t \geq 0\}$.

Let us now return to our $F/G/k$ queueing system.

Theorem 3:

If G is nonlattice, either G or F is continuous, μ_G and μ_F are finite and satisfy $\mu_G > k\mu_F$, then G and F are identifiable from $\{C(t), t \geq 0\}$.

Proof:

Since $\mu_G > k\mu_F$ it follows that with probability one the queue size will converge to infinity, and hence after some finite amount of time the output process will behave as the superposition of k independent renewal processes each having interarrival distribution G . Our continuity assumption assures that an arrival and departure will not occur simultaneously and hence departure epochs will correspond one-to-one with points of decrease of $C(t)$. By Theorem 2

it follows that G is identifiable. Since points of increase of $C(t)$ correspond in a one-to-one manner with arrival epochs it follows that F is identifiable by the empirical c.d.f. Q.E.D.

Remark:

As in the remark following Theorem 1, we note that if the data consists of the arrival and departure epochs then the continuity assumption may be dropped.

REFERENCES

- [1] Brown, M., "An $M/G/\infty$ Estimation Problem," To Appear Ann. of Math. Statist., (1969).
- [2] Cox, D.R., "Some Problems of Statistical Analysis Connected with Congestion," PROCEEDINGS OF THE SYMPOSIUM ON CONGESTION THEORY, University of North Carolina Press, Chapel Hill, (1965).
- [3] Kendall, D.G. and T. Lewis, "On the Structural Information Contained in the Output of $G.I./G/\infty$," Z. Wahrscheinlichkeitstheorie Verw., Vol. 4, pp. 144-148, (1965).
- [4] Kiefer, J. and J. Wolfowitz, "On the Theory of Queues with Many Servers," Trans. Amer. Math. Soc., Vol. 78, pp. 1-18, (1955).
- [5] Kiefer, J. and J. Wolfowitz, "On the Characteristics of the General Queueing Process, with Applications to Random Walk," Ann. Math. Statist., Vol. 27, pp. 147-161, (1956).

Unclassified

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

ORIGINATING ACTIVITY (Corporate author)

University of California, Berkeley

2a. REPORT SECURITY CLASSIFICATION

Unclassified

2b. GROUP

REPORT TITLE

IDENTIFIABILITY IN GI/G/k QUEUES

DESCRIPTIVE NOTES (Type of report and, inclusive dates)

Research Report

AUTHOR(S) (First name, middle initial, last name)

Sheldon M. Ross

REPORT DATE

August 1970

7a. TOTAL NO. OF PAGES

10

7b. NO. OF LINES

5

1. CONTRACT OR GRANT NO.

DA-31-124-ARO-D-331

2. PROJECT NO.

20014501B14C

9a. ORIGINATOR'S REPORT NUMBER(S)

ORC 70-26

9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)

3. DISTRIBUTION STATEMENT

This document has been approved for public release and sale; its distribution is unlimited.

4. SUPPLEMENTARY NOTES

NONE

12. SPONSORING MILITARY ACTIVITY

U.S. Army Research Office-Durham
Box CM, Duke Station
Durham, North Carolina 27706

5. ABSTRACT

SEE ABSTRACT.

Unclassified

Security Classification

Unclassified

Security Classification

14

KEY WORDS

LINK A

LINK B

LINK C

ROLE

WT

ROLE

WT

ROLE

WT

GI/G/k Queues

Identifiability

Superposition of Renewal Processes

Estimate

Stationary Process

Ergodic Theorem

DD FORM 1 NOV 66 1473 (BACK)

DN 0101-007-6021

Unclassified

Security Classification

A-31409