AD-756 668

BOUNDS ON THE DELAY DISTRIBUTION IN GI/G/1 QUEUES

Sheldon M. Ross

California University

Prepared for:

Office of Naval Research
Army Research Office-Durham

January 1973

ORC 73-1
JANUARY 1973

BOUNDS ON THE DELAY DISTRIBUTION IN GI/G/1 QUEUES

by

SHELDON M. ROSS

OPERATIONS RESEARCH CENTER

COLLEGE OF ENGINEERING
UNIVERSITY OF CALIFORNIA · BERKELEY

R

BOUNDS ON THE DELAY DISTRIBUTION IN GI/G/1 QUEUES

by

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

JANUARY 1973 ORC 73-1

This research has been partially supported by the U.S. Army Research Office-Durham under Contract DA-31-124-ARO-D-331 and the Office of Naval Research under Contract N00014-69-A-0200-1036 with the University of California. Reproduction in whole or in part is permitted for any purpose of the United States Government.

iq

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)

1. ORIGINATING ACTIVITY (Corporate author) University of California, Berkeley	2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP

3. REPORT TITLE

BOUNDS ON THE DELAY DISTRIBUTION IN GI/G/1 QUEUES

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

Research Report

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

Sheldon M. Ross

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
January 1973	~~1~~ 12	4

8a. CONTRACT OR GRANT NO. DA-31-124-ARO-D-331 b. PROJECT NO. 20014501B14C c. d.	9a. ORIGINATOR'S REPORT NUMBER(S) ORC 73-1 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)

10. DISTRIBUTION STATEMENT

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

11. SUPPLEMENTARY NOTES Also supported by the Office of Naval Research under Contract N00014-69-A-0200-1036.	12. SPONSORING MILITARY ACTIVITY U.S. Army Research Office-Durham Box CM, Duke Station Durham, North Carolina 27706

13. ABSTRACT

SEE ABSTRACT.

Unclassified

Security Classification

A-31408

Unclassified
Security Classification

14 KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Delay Distribution						
GI/G/1 Queues						
Martingale						
Stopping Time						
NBU						
NWU						
IFR						
DFR						

ABSTRACT

Bounds are obtained for the limiting distribution of the delay in queue for a GI/G/1 system via Martingale theory. These bounds are somewhat stronger than similar bounds recently obtained by Kingman. Simplifications of the bounds are obtained in the special cases where the service distribution is either IFR, DFR, NBU or NWU.

1. INTRODUCTION

Consider the usual GI/G/1 queue with interarrival times between customers $X_1, X_2, \ldots$ and service times $Y_1, Y_2, \ldots,$ where $E(Y_i) < E(X_i) < \infty$. Let $U_i = Y_i - X_i$ and assume that there exists a nonzero value θ such that $E\left[e^{\theta U_i}\right] = 1$. If such a value of θ exists then by Jensen's inequality it must be positive since $EU_i < 0$. Let D_n denote the delay in queue of the n^{th} customer, and let

$$\bar{D}(t) = \lim_{n\to\infty} P\{D_n > t\} .$$

In [1] and [2] the following inequality was proven by Kingman

$$ae^{-\theta t} \leq \bar{D}(t) \leq e^{-\theta t}$$

where

$$a = \inf_{t>0}\left[\int_t^\infty dF(y) \Big/ \int_t^\infty e^{\theta(y-t)}dF(y)\right]$$

and where F is the distribution of U_i. Kingman proved the right side of the above inequality in [1] by using Kolmogorov's inequality for Martingales, and used a different technique in [2] to obtain the complete inequality. Following the Martingale approach of Kingman but using an appropriate stopping time rather than Kolmogorov's inequality a somewhat sharper inequality will now be obtained.

2. THE INEQUALITIES

As was shown by Lindley [3]

$$\bar{D}(t) = P\left\{\sum_1^n U_i > t \quad \text{for some} \quad n = 1,2,\ \ldots\right\}.$$

If we let $Z_n = e^{\theta \sum_1^n U_i}$, then as Z_n is the product of independent random variables each with mean 1, it follows that $\left\{Z_n\ n \geq 1\right\}$ is a Martingale. For a fixed positive constant A, define the stopping time N_A by

$$N_A = \text{1st} \quad n \quad \text{such that either} \quad Z_n > e^{\theta t} \quad \text{or} \quad Z_n < e^{-\theta A}.$$

As is well known the moment generating function of N_A exists in a region about 0 and thus by Martingale theory

$$1 = E(Z_n) = E\left(Z_{N_A}\right)$$

$$(1) \qquad = E\left[e^{\theta \sum_1^{N_A} U_i} \,\middle|\, \sum_1^{N_A} U_i > t\right] P\left(\sum_1^{N_A} U_i > t\right) + E\left[e^{\theta \sum_1^{N_A} U_i} \,\middle|\, \sum_1^{N_A} U_i < -A\right] P\left(\sum_1^{N_A} U_i < -A\right).$$

Now,

$$(2) \qquad E\left[e^{\theta \sum_1^{N_A} U_i} \,\middle|\, \sum_1^{N_A} U_i > t\right] = e^{\theta t} E\left[e^{\theta\left(\sum_1^{N_A} U_i - t\right)} \,\middle|\, \sum_1^{N_A} U_i > t\right].$$

By conditioning on N_A and $\sum_1^{N_A-1} U_i$ we obtain that

$$(3)\quad \inf_{0\le r} E\left[e^{\theta(U_1-r)} \mid U_1 > r\right] \le E\left[e^{\theta\left(\sum_1^{N_A} U_i - t\right)} \mid \sum_1^{N_A} U_i > t\right] \le \sup_{0\le r} E\left[e^{\theta(U_1-r)} \mid U_1 > r\right].$$

Also, since

$$\lim_{A\to\infty} P\left(\sum_1^{N_A} U_i > t\right) = P\left\{\sum_1^{n} U_i > t \quad \text{for some} \quad n < \infty\right\} = \bar{D}(t)$$

and

$$\lim_{A\to\infty} E\left[e^{\theta \sum_1^{N_A} U_i} \mid \sum_1^{N_A} U_i < -A\right] = 0$$

we obtain from (1), (2), and (3), by letting $A \to \infty$ that

$$(4)\quad \frac{e^{-\theta t}}{\sup_{0\le r} E\left[e^{\theta(U_1-r)} \mid U_1 > r\right]} \le \bar{D}(t) \le \frac{e^{-\theta t}}{\inf_{0\le r} E\left[e^{\theta(U_1-r)} \mid U_1 > r\right]}.$$

Note that the left side inequality is just the left side of Kingman's inequality while the right sided inequality is stronger than Kingman's.†

A somewhat weaker though probably more useful inequality based on the service distribution can be obtained from the above as follows:

$$E\left[e^{\theta(U_1-r)} \mid U_1 > r\right] = E\left[e^{\theta[Y_1-(X_1+r)]} \mid Y_1 > X_1 + r\right].$$

Hence, by conditioning on X_1 we obtain that

$$\inf_{0\le s} E\left[e^{\theta(Y_1-s)} \mid Y_1 > s\right] \le E\left[e^{\theta(U_1-r)} \mid U_1 > r\right] \le \sup_{0\le s} E\left[e^{\theta(Y_1-s)} \mid Y_1 > s\right]$$

---

†The sup and inf are taken over all nonnegative values of r for which $P(U_1 > r) > 0$.

As this is true for all r, we obtain that

$$\sup_{0\le r} E\left[e^{\theta(U_1-r)} \mid U_1 > r\right] \le \sup_{0\le s} E\left[e^{\theta(Y_1-s)} \mid Y_1 > s\right]$$

$$\inf_{0\le r} E\left[e^{\theta(U_1-r)} \mid U_1 > r\right] \ge \inf_{0\le s} E\left[e^{\theta(Y_1-s)} \mid Y_1 > s\right] .$$

Thus, from (4) we obtain[†]

$$(5) \qquad \frac{e^{-\theta t}}{\sup\limits_{0\le s} E\left[e^{\theta(Y_1-s)} \mid Y_1 > s\right]} \le \bar{D}(t) \le \frac{e^{-\theta t}}{\inf\limits_{0\le s} E\left[e^{\theta(Y_1-s)} \mid Y_1 > s\right]}$$

A very important special case occurs when the service times are exponentially distributed with mean $1/\mu$. That is, when the system is a G/M/1 queue. In this case, using the lack of memory of the exponential distribution we have that the conditional distribution of $Y_1 - s$ given that $Y_1 > s$ is just exponential with mean $1/\mu$. Hence,

$$\sup_{s} E\left[e^{\theta(Y_1-s)} \mid Y_1 > s\right] = \inf_{s} E\left[e^{\theta(Y_1-s)} \mid Y_1 > s\right] = \frac{\mu}{\mu - \theta}$$

and thus in the G/M/1 case

$$\bar{D}(t) = \frac{\theta}{\mu - \theta} e^{-\theta t} .$$

In certain special cases the sup and inf in equation (5) can be more easily expressed. We say that the service distribution G is NBU if

$$\bar{G}(s + t) \le \bar{G}(s)\bar{G}(t) \qquad \text{for all } s,t \ge 0$$

and it is said to be NWU if

[†]The sup and inf in equation (5) are taken over all nonnegative values of s for which $P(Y_1 > s) > 0$.

$$\bar{G}(s + t) \geq \bar{G}(s)\bar{G}(t) \quad \text{for all} \quad s,t \geq 0$$

where $\bar{G}(t) = 1 - G(t)$. Thus G is NBU (NWU) means that the remaining service time of a customer who has already been in service for some fixed time is always stochastically smaller (larger) than the service time of a customer just entering service.†

Since V being stochastically larger than W implies that $E[f(V)] \geq E[f(W)]$ for all increasing functions f , we obtain that if G is NBU with $G(0) = 0$ then

$$\sup_{0 \leq s} E\left[e^{\theta(Y_1 - s)} \mid Y_1 > s\right] = E\left[e^{\theta Y_1}\right]$$

while if G is NWU with $G(0) = 0$ then

$$\inf_{0 \leq s} E\left[e^{\theta(Y_1 - s)} \mid Y_1 > s\right] = E\left[e^{\theta Y_1}\right] .$$

If we make the stronger (than NBU) assumption that G is IFR (that is, that $\dfrac{\bar{G}(s + t)}{\bar{G}(t)}$ decreases in t for all s) then it easily follows when $G(0) = 0$ that

$$\sup_{0 \leq s} E\left[e^{\theta(Y_1 - s)} \mid Y_1 > s\right] = E\left[e^{\theta Y_1}\right]$$

and

$$\inf_{0 \leq s} E\left[e^{\theta(Y_1 - s)} \mid Y_1 > s\right] = \lim_{s \to M} E\left[e^{\theta(Y_1 - s)} \mid Y_1 > s\right]$$

where $M = \sup\{t : \bar{G}(t) > 0\}$.

---

†The terminology NBU (new better than used) originated in reliability literature where it means that if G is the distribution of the lifetime of a component then the remaining life of any s year old (that is, any used) item is stochastically smaller than the lifetime of a new item.

Similarly if G is assumed to be DFR (that is, if it is assumed that $\frac{\bar{G}(s+t)}{\bar{G}(t)}$ increases in t for all s) then it follows when $G(0) = 0$ that

$$\sup_{0 \le s} E\left[e^{\theta(Y_1 - s)} \mid Y_1 > s\right] = \lim_{s \to \infty} E\left[e^{\theta(Y_1 - s)} \mid Y_1 > s\right]$$

$$\inf_{0 \le s} E\left[e^{\theta(Y_1 - s)} \mid Y_1 > s\right] = E\left[e^{\theta Y_1}\right] .$$

The importance of DFR service time distributions partly derives from the fact that if the actual service distribution is a mixture of other distributions each of which is DFR (for instance, each may be exponential) then the service distribution is also DFR .

REFERENCES

[1] Kingman, J. F. C., "A Martingale Inequality in the Theory of Queues," Proceedings Cambridge Philosophical Society, 59, 359-361 (1964).

[2] Kingman, J. F. C., "Inequalities in the Theory of Queues," Royal Statistical Society Journal, Vol. 32, 102-110 (1970).

[3] Lindley, D. V., Proceedings Cambridge Philosophical Society, 48, 277-289 (1952).

[4] Marshall, K. T., "Some Inequalities in Queueing," Operations Research, 16, 651-665 (1968).