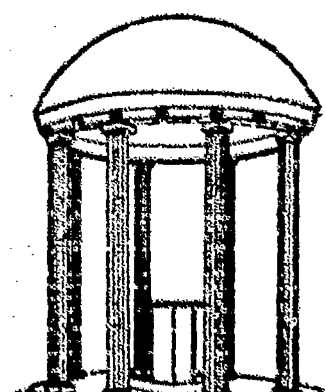
See 1473 (1)

F.B.

OPERATIONS RESEARCH AND SYSTEMS ANALYSIS

UNIVERSITY OF NORTH CAROLINA
AT CHAPEL HILL



Maximum Likelihood Estimation of the Distribution of the Sum of Three Independent Exponential Random Variables

George S. Fishman

Technical Report No. 76-7
May, 1976

Curriculum in Operations Research
and Systems Analysis

University of North Carolina at Chapel Hill

This research was supported by the Office of Naval Research under contract N00014-67-A-[illegible]. Reproduction in whole or in part is permitted for any purpose of the United States government.

ABSTRACT

This paper describes a procedure for computing the maximum likelihood estimates of the parameters of the distribution of the sum of three independent exponential random variables. By fitting sample interevent time data from a real system to this distribution, one can create a simulation of the system that exploits the regenerative representation of queueing systems [3] to analyze the simulation's output by relatively elementary statistical methods. The paper also describes computation of the sample asymptotic covariance matrix and an implementation of the likelihood ratio for testing six hypotheses that are special cases of interest. A set of FORTRAN subroutines for executing these procedures appears in the Appendix.

1. Introduction

The purpose of this paper is to describe a procedure for computing the maximum likelihood estimates (MLE) of the parameters of the distribution of the sum of three independent exponentially distributed random variables. Although the case of equal parameters yields an Erlang distribution, for which the MLE are known, the more general case has received little attention in the statistical literature. Two possible reasons for this omission occur to the writer. Firstly, since the corresponding MLE equations are not amenable to analytical solution, one needs to employ numerical analytic techniques to solve them. Conceptually, the presence of multiple maxima makes this an onerous approach. Secondly, since the distribution has three parameters, the principle of parsimony encourages one to use alternative two parameter distributions whenever a fit of equal or almost equal quality can be obtained. These distributions include the gamma, lognormal and Weibull. Choi and Wette [1] describe a procedure for computing the gamma MLE. Thoman, Bain and Antle [13] describe a procedure for computing the Weibull MLE. Although both procedures rely on the Newton-Raphson iterative method no unusual problems arise. For the lognormal distribution the MLE relate directly to the MLE for the corresponding normal distribution. Johnson and Kotz [9] discuss issues related to the MLE for these distributions, including bias removal.

Given the attractions of alternative distributions, a relatively strong justification for pursuing the research presented here seems in order. Recent developments in the field of discrete event simulation provide this justification. In [5,6] Fishman points out that in the simulation of queueing systems one could use the exit of the system from the empty and idle state to demarcate the sample path of a stochastic process of interest into independent segments each of which obeys the same probability law. This demarcation enables one to use relatively elementary statistical methods to compute point and interval estimates for population parameters of interest [5,6]. The most appealing theoretical feature of this observation is that the i.i.d. property holds regardless of the distributions of interarrival and service times. The most unappealing feature arises when either the activity level increases or the number of servers increases for a given activity level. In particular, the frequency with which the system exits the empty and idle state declines dramatically. In turn, this can result in excessively long simulation runs if one is determined to collect a prespecified number of i.i.d. segments.

In [2] and [3] Crane and Iglehart introduce the more general notion of a regenerative process into the analysis of simulation output. In particular, any state can serve as a demarcating state, provided that statistical behavior after entry into that state is independent of behavior prior to entry and that the state occurs infinitely often. States with these properties are called regenerative. If one can identify all such states then one can use the most frequently occurring one to demarcate the specified number of i.i.d. segments. If the interevent distributions are exponential then all states can serve this demarcating purpose. Since exponentiality is too restrictive an assumption in general, Crane and Iglehart [4] attempt to identify approximate regenerative states. Their procedure calls for a careful scrutiny of the particular system being simulated.

An alternative approach to realizing the regenerative property arises when interevent times have continuous unimodal distributions. Then a theoretical basis exists for approximating each of these distributions by the distribution of the sum of an arbitrary number of independent exponential random variables. In particular, one way to look at this is to consider the polynomial approximation to the corresponding characteristic function where the reciprocals of the roots of the polynomial, which are real for unimodality, are the means of the exponential random variables.† If one adopts this characterization then interevent times in the simulation become sums of independent exponential random variables. Suppose, interarrival times are representable as the sum of two independent exponential random variables and service times are exponential. Then by adding a new entry to the state vector that characterizes which of the two stages the next arrival occupies, one provides the mechanism for realizing regenerative states. If service times are representable as the sum of three independent exponential random variables then three additional entries in the state vector to keep track of the number of jobs in each stage enable one to exploit the regenerative property again. The price paid for this ability is the increased bookkeeping for the state vector, an efficient approach to which is described in [7].

Although the foregoing discussion motivates the use of distributions of sums of independent exponentials, a procedure for implementing the approach is practice remains to be developed. Ideally, one would like to fit such a distribution by the distribution of the sum of a large number of exponential variates and, through a formal hypothesis testing procedure, reduce that sum to the minimal number necessary to

†This assumes a polynomial in $i\omega$ where $i = \sqrt{-1}$.

account for variation in the data. The present paper describes a first step in this direction in Section 2 by fitting the sum of three independent exponential random variables and then testing six hypotheses designed to reduce the length of the state vector. In particular, Section 2 describes a procedure for finding the MLE,their sample asymptotic covariance matrix and for using the likelihood ratio to test hypotheses. The steps outlined in Section 2 are implemented in a set of FORTRAN subroutines in the Appendix.

2. The Procedure

Let Y_1, Y_2 and Y_3 be independent random variables from $E(a)$, $E(b)$ and $E(c)$, respectively, where $E(\theta)$ denotes the exponential distribution

$$(1) \qquad f(x) = \begin{cases} e^{-x/\theta}/\theta & 0 \le x \le \infty \quad 0 < \theta \\ 0 & \text{elsewhere.} \end{cases}$$

Then $X = Y_1 + Y_2 + Y_3$ has the probability density function (p.d.f.)

$$(2) \qquad f(x,a,b,c) = g(x,a,b,c) + g(x,b,a,c) + g(x,c,a,b)$$

where

$$(3) \qquad g(x,\theta,\phi,\rho) = \theta e^{-x/\theta}/(\theta-\phi)(\theta-\rho).$$

Given a sample $X_1,\ldots,X_n$ from (2), we wish to compute $\hat{a}$, $\hat{b}$, $\hat{c}$, the MLE of a, b and c, respectively. These follow from maximization of the likelihood function

$$(4) \qquad L = \prod_{i=1}^{n} f(X_i,a,b,c) .$$

Here $\hat{a}$, $\hat{b}$, $\hat{c}$ asymptotically have the trivariate normal distribution with means a, b, and c, respectively, and the minimum variance covariance matrix $\underline{\Sigma}$, where [10]

$$(5) \qquad \underset{\sim}{\Sigma} = \begin{bmatrix} E\left(\frac{\partial \ln L}{\partial a}\right)^2 & E\left(\frac{\partial \ln L}{\partial a}\,\frac{\partial \ln L}{\partial b}\right) & E\left(\frac{\partial \ln L}{\partial a}\,\frac{\partial \ln}{\partial c}\right) \\ & E\left[\frac{\partial \ln L}{\partial b}\right]^2 & \left(\frac{\partial \ln L}{\partial b}\,\frac{\partial \ln L}{\partial c}\right) \\ & & E\left(\frac{\partial \ln L}{\partial c}\right)^2 \end{bmatrix}^{-1} .$$

To obtain the MLE one usually solves

$$(6) \qquad \frac{\partial \ln L}{\partial \theta} = \sum_{i=1}^{n} \frac{1}{f(X_i,a,b,c)} \cdot \frac{\partial f(X_i,a,b,c)}{\partial \theta} = 0 \qquad \theta = a,b,c$$

simultaneously for a, b and c. In the present case (6) does not admit an analytical solution. Moreover, the only sufficient statistics appear to be $X_1,\ldots,X_n$ which do little to ease the computational burden of a numerical solution.

Feasible Region

Although the possibility of multiple maxima makes maximization of L difficult in general, we can reduce some of this difficulty by noting that

$$\begin{aligned} (7) \qquad f(x,a,b,c) &= f(x,a,c,b) = f(x,b,a,c) \\ &= f(x,b,c,a) = f(x,c,a,b) \\ &= f(x,c,b,a) . \end{aligned}$$

This implies that L has at least 6 maxima of equal magnitude. Introducing the constraints

(8) $$a \le b \le c$$

removes this ambiguity. One can also show that

(9) $$a^2 \frac{\partial \ln L}{\partial a} + b^2 \frac{\partial \ln L}{\partial b} + c^2 \frac{\partial \ln L}{\partial c} = 0$$

leads to

(10) $$\hat{a} + \hat{b} + \hat{c} = \bar{X}$$

$$\bar{X} = (1/n) \sum_{i=1}^{n} X_i \ .$$

Now the constraints (8) and (10) imply

(11) $$0 \le 2\hat{a} \le \bar{X} - \hat{c}$$

$$\bar{X} - \hat{a} \le 2c \le \bar{X}$$

which define the feasible region in the $\hat{a} - \hat{c}$ space of Figure 1 where maximization of L is to occur.

The arcs and nodes of the feasible region in Figure 1 play a special role here. In particular, arcs AB, BC and AC correspond to hypotheses 1, 2 and 3 in Table 1 and nodes B, A and C correspond to the Erlang hypotheses 4, 5 and 6. In addition to examining these special cases in the process of maximization of L, one can use the likelihood ratio test to evaluate the effect of assuming that one of these special cases represents the underlying structure of f in (1). This issue is discussed shortly.

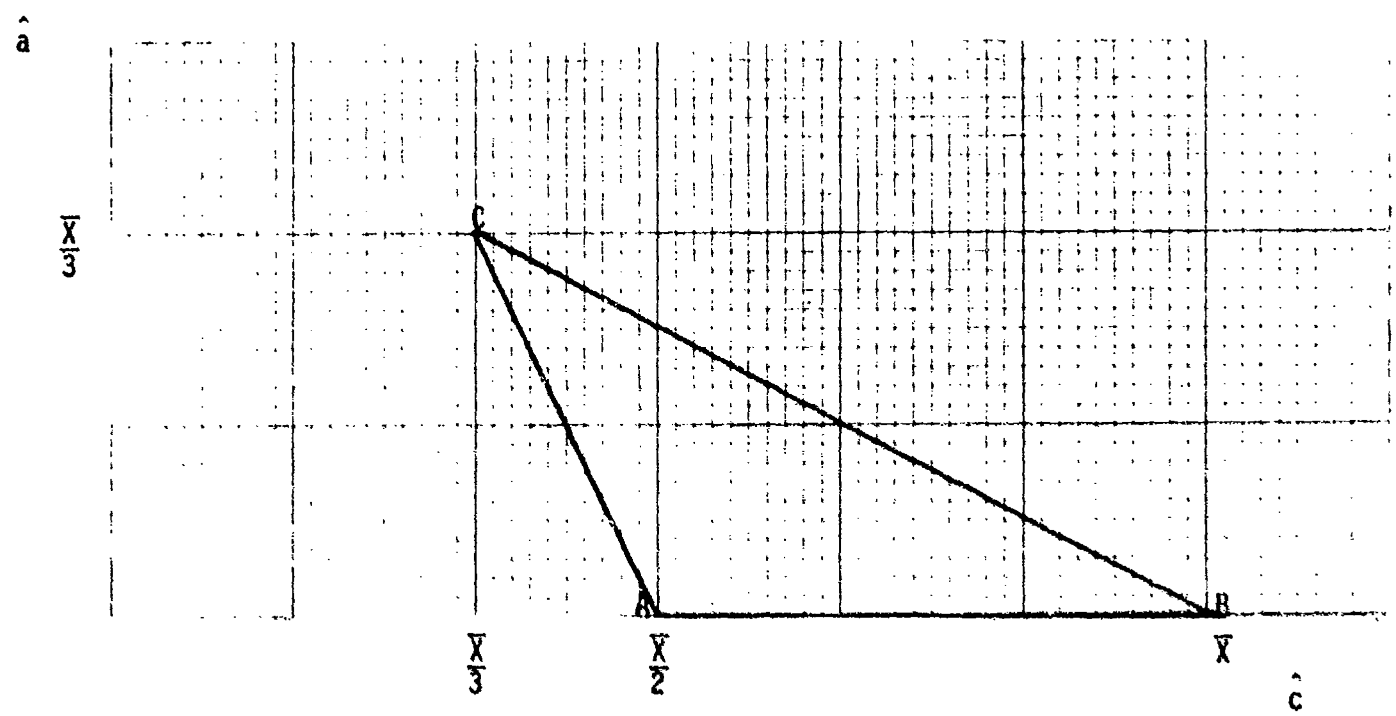


Figure 1 Feasible Region for MLE

<u>Computational Considerations</u>

The search for a maximum for L has now been restricted to the triangle in Figure 1. The set of FORTRAN subprogram listed in the Appendix effects a grid search on $\hat{a}$ in user specified increments of δ over $[0, \bar{X}/3]$ and for each $\hat{a}$ performs a binary search for $\hat{c}$ in

Table 1

Distributions and Derivatives Under Alternative Hypotheses

i	H_i	$f(x,a,b,c)$	$\frac{\partial f(x,a,b,c)}{\partial a}$	$\frac{\partial f(x,a,b,c)}{\partial b}$	$\frac{\partial f(x,a,b,c)}{\partial c}$
0	$a \le b \le c$	$g(x,a,b,c)+g(x,b,a,c)+g(x,c,a,b)$	$h_1(x,a,b,c)$	$h_1(x,b,a,c)$	$h_1(x,c,a,b)$
1	$a=0, b \le c$	$g(x,b,c,0)+g(x,c,b,0)$	-	$h_1(x,b,c,0)$	$h_1(x,c,b,0)$
2	$a=b \le c$	$g(x,a,c,c)[x(a-c)-ac]/a^2+g(x,c,a,a)$	$h_3(x,c,a)$	-	$h_2(x,a,c)$
3	$a \le b=c$	$g(x,c,a,a)[x(c-a)-ac]/c^2+g(x,a,c,c)$	$h_2(x,c,a)$	-	$h_3(x,a,c)$
4	$a=b=0<c$	$g(x,c,0,0)$	-	-	$g(x,c,0,0)(x/c-1)/c$
5	$a=0, b=c$	$xg(x,c,0,0)/c$	-	-	$xg(x,c,0,0)(x/c-2)/c^2$
6	$a=b=c$	$x^2g(x,c,0,0)/2c^2$	-	-	$x^2g(x,c,0,0)(x/c-3)/2c^3$

$$h_1(x,\theta,\phi,\rho)=g(x,\theta,\phi,\rho)[1/\theta+x/\theta^2-1/(\theta-\phi)-1/(\theta-\rho)]+g(x,\phi,\theta,\rho)/(\phi-\theta)+g(x,\rho,\theta,\phi)/(\rho-\theta)$$

$$h_2(x,\theta,\rho)=g(x,\rho,\theta,\theta)[1/\rho+x/\rho^2-2/(\rho-\theta)]+g(x,\theta,\rho,\rho)[(\rho-\theta)x+\theta(\theta+\rho)]/(\rho-\theta)\theta^2$$

$$h_3(x,\theta,\rho)=g(x,\rho,\theta,\theta)(\{[(\rho-\theta)x-\theta\rho][x/\rho^2-1/\rho-2/(\rho-\theta)]+x-\theta)/\rho^2+2g(x,\theta,\rho,\rho)/(\theta-\rho)$$

$[(\overline{X}-\hat{a})/2, \overline{X}-2\hat{a}]$ to within the tolerance δ. The search for $\hat{c}$ solves $\partial \ell n\, L/\partial c = 0$. Expression (10) yields $\hat{b}$ and the search for the maximum is effected by computation and comparison of $\ell n\, L$ for each computed set of $\hat{a}$, $\hat{b}$ and $\hat{c}$. Since substantial experience with the UPDATE subroutine using a complete grid search in ESTIMA failed to reveal more than one maximum, ESTIMA was modified to terminate the search once a maximum has been found.

The ARC and NODE subroutines enable one to check the arcs AB, BC and AC and the nodes A, B and C for solutions that might give improvement. Also HYP123 and ERLANG use the results of ARC and NODE, respectively, to test the hypotheses in Table 1.

Computation of Covariance Matrix

The estimation of the covariance matrix under H_0, H_1, H_2 and H_3 uses

$$(12) \qquad E\left(\frac{\partial\, \ell n\, L}{\partial \theta}\ \frac{\partial\, \ell n\, L}{\partial \phi}\right) = n\int_0^\infty \frac{\partial f(x,a,b,c)}{\partial \theta}\cdot\frac{\partial f(x,a,b,c)}{\partial \phi}\cdot\frac{1}{f(x,a,b,c)}\,dx$$

together with the expressions in Table 1 in ESTIMA and HYP123. These subroutines employ numerical integration, as described in [12, p.923] to evaluate $\underline{\Sigma}$, using $\hat{a}$, $\hat{b}$, $\hat{c}$ in place of a, b and c respectively. Although the weights in the W and Y arrays apply for double precision computation, experience has shown little loss of accuracy by using single precision. Figure 2 offers an example of the output for 100 observations drawn from f(x,1,5,12).

Likelihood Ratio Test

Since parsimony clearly has advantages in modeling, one wants to test the hypotheses in Table 1 to see if one or two parameters can be eliminated from the representation (1). Let $L(\underline{X}, \hat{a}_i, \hat{b}_i, \hat{c}_i)$ denote the maximum of the likelihood function under H_i where $i = 0$ corresponds to (1) and $\underline{X} = (X_1, \ldots, X_n)$. For example, $L(\underline{X}, \hat{a}_1, \hat{b}_1, \hat{c}_1) = L(\underline{X}, 0, \hat{b}_1, \hat{c}_1)$ and $L(\underline{X}, \hat{a}_2, \hat{b}_2, \hat{c}_2) = L(\underline{X}, \hat{a}_2, \hat{a}_2, \hat{c}_2)$. Then the likelihood ratio

$$(13) \qquad R_i = L(\underline{X}, \hat{a}_i, \hat{b}_i, \hat{c}_i)/L(\underline{X}, \hat{a}_0, \hat{b}_0, \hat{c}_0) \qquad i = 1, \ldots, 6$$

lies in $(0,1)$. The closer R_i is to unity the more credible is the hypothesis. Although the distribution of R_i under H_i is beyond our reach it is known that as n increases the distribution of $-2 \ln R_i$ converges to the chi-square distribution with degrees of freedom equal to the number of constraints imposed by the hypothesis [10]. For H_1, H_2 and H_3 there is 1 degree of freedom; for H_4, H_5 and H_6, there are 2 degrees of freedom. Therefore

$$(14) \qquad \mathrm{pr}(R_i \geq e^{-\chi_f^2(1-\alpha)/2}) = 1 - \alpha$$

where $\chi_f^2(1-\alpha)$ denotes the $1 - \alpha$ critical value for f degrees of freedom and $f = 1$ for $i = 1, \ldots, 3$ and $f = 2$ for $i = 4, \ldots, 6$. Table 2 shows critical values of R_i corresponding to tests of selected sizes.

MAXIMUM LIKELIHOOD ESTIMATION

F(X)=G(X,A,B,C)+G(X,B,A,C)+G(X,C,A,B)

G(X,T,P,R)=T*EXP(-X/T)/((T-P)*(T-R))

N= 100 SAMPLE MEAN= 0.209147E 02 SAMPLE VARIANCE= 0.250988E 03

DELTA= 0.697155E-02

A= 0.874065E 00 B= 0.558364E 01 C= 0.148639E 02

COVARIANCE MATRIX

0.137369E 01	-0.359012E 01	0.304326E 01
	0.195890E 02	-0.222481E 02
		0.316986E 02

CORRELATION MATRIX

0.100000E 01	-0.692083E 00	0.461184E 00
	0.100000E 01	-0.892826E 00
		0.100000E 01

HYPOTHESIS 1: A=0, B<=C

B= 0.646436E 01 C= 0.144503E 02

VAR(B)= 0.170784E 02 VAR(C)= 0.361412E 02 COV(B,C)=-0.231085E 02

CORR(B,C)=-0.930136E 00

LIKELIHOOD RATIO= 0.781658E 00

HYPOTHESIS 2: A=B<=C

B= 0.256668E 01 C= 0.157813E 02

VAR(B)= 0.575024E 00 VAR(C)= 0.102919E 02 COV(B,C)=-0.171543E 01

CORR(B,C)=-0.705151E 00

LIKELIHOOD RATIO= 0.722563E 00

```
HYPOTHESIS 3:    A<=B=C

                  A= 0.272369E-01      C= 0.104437E 02

     VAR(A)= 0.255362E 00     VAR(C)= 0.950815E 00     COV(A,C)=-0.172593E 00

                         CORR(A,C)=-0.349580E 00

                      LIKELIHOOD RATIO= 0.575515E 00


HYPOTHESIS 4:    A=B=0

                            C= 0.209147E 02

    .95 LOWER POINT= 0.173517E 02      .95 UPPER POINT= 0.257062E 02

                      LIKELIHOOD RATIO= 0.198047E-04


HYPOTHESIS 5:    A=0, B=C

                            C= 0.104573E 02

    .95 LOWER POINT= 0.914665E 01      .95 UPPER POINT= 0.120730E 02

                      LIKELIHOOD RATIO= 0.563763E 00


HYPOTHESIS 6:    A=B=C

                            C= 0.697155E 01

    .95 LOWER POINT= 0.620519E 01      .95 UPPER POINT= 0.783313E 01

                      LIKELIHOOD RATIO= 0.676001E-03
```

Figure 2 (continued)

Table 2

Critical Values of R_i for Tests of Selected Sizes

d.f. \ α	0.01	0.025	0.05	0.10	0.25
1	0.9999	0.9995	0.9983	0.9921	0.9505
2	.9900	.9750	.9500	.9000	.7500

In the case of 2 d.f. the chi square distribution is $E(1)$. Therefore, R_i^2 is the probability that under H_i $(i=4,5,6)$ one would observe a likelihood ratio less than† R_i. For example, under H_5 in Figure 2 $R_5 = 0.5658$ and $R_5^2 = 0.3201$.

Confidence Intervals

Let us first concentrate on H_4, H_5 and H_6. Under H_i $\hat{c}_i = X/(i-3)$ and $n\,\hat{c}_i/c$ has the chi-square distribution with $(i-3)n/2$ degrees of freedom. The ERLANG subroutine uses this fact to compute a confidence interval for c and relies on the CHISQ subroutine to provide critical values of chi-square.

For H_0, H_1, H_2 and H_3 no similar theory is available. However, if n is sufficiently large, one can compute approximate individual confidence intervals for a, b and c, using the estimated variances in the corresponding covariance matrix. Experience with the set of subprograms in the appendix has revealed that even for $n \sim 100$ the sample $\text{var}(\hat{a})$, $\text{var}(\hat{b})$, $\text{var}(\hat{c})$ are large relative to $\hat{a}$, $\hat{b}$ and $\hat{c}$ respectively.

†Here R_i^2 is called the P-value. See [8].

Bias Considerations

In small and moderate size samples $\hat{a}$, $\hat{b}$ and $\hat{c}$ are biased. In particular, experience has shown that $\hat{a}$ overestimates a and $\hat{c}$ underestimates c. Since $\hat{c}$ does most to affect the shape of the tail of the distribution we especially want to consider ways of reducing bias for this quantity. One approach to bias reduction uses the *jackknife* method [11].

The elementary form of the jackknife method removes bias to order $1/n$. Suppose $\hat{c}$ is computed using n observations and $\hat{c}^{(1)}$ and $\hat{c}^{(2)}$ are computed using the first $m = n/2$ observations and the last $m = n/2$ observations respectively. Then one can easily show that

$$(15) \qquad \tilde{c} = 2\hat{c} - (\hat{c}^{(1)} + \hat{c}^{(2)})/2$$

is free from bias to order $1/n$. Notice that the computation of $\tilde{c}$ requires 3 passes through the estimation procedure.

More powerful jackknife methods of bias reduction are available [11]. Our reluctance to incorporate any one of them into the estimation procedure is a consequence of the additional cost they imply. However, a user of the estimation procedure in the Appendix can easily write a bias reduction program to use in conjunction with ESTIMA.

3. References

1. Choi, S.C. and R. Wette, "Maximum Likelihood Estimation of the Parameters of the Gamma Distribution and Their Bias," *Technometrics*, Vol. 11, No. 4, November 1969, pp. 623-690.

2. Crane, G.S., "Estimation in Multiserver Queueing Simulations," *Operations Research*, Vol. 22, No. 1, January-February, 1974, pp. 72-78.

3. Crane, M.A. and D. Iglehart, "Simulating Stable Stochastic Systems, III: Regenerative Processes and Discrete-Event Simulations," *Operations Research*, Vol. 23, No. 1, January-February, 1975, pp. 33-45.

4. Crane, M.A. and D. Iglehart, "Simulating Stable Stochastic Systems IV: Approximation Techniques," *Management Science*, Vol. 21, 1975, pp. 1215-1224.

5. Fishman, G.S., "Statistical Analysis for Queueing Simulations," *Management Science*, Vol. 20, No. 3, November 1973, pp. 363-369.

6. Fishman, G.S., "Estimation in Multiserver Queueing Simulations," *Operations Research*, Vol. 22, No. 1, January-February, 1974, pp. 72-78.

7. Fishman, G.S., *Discrete Event Simulation*, to appear.

8. Gibbons, J.D. and J.W. Pratt, "P-Values: Interpretation and Methodology," *The American Statistician*, Vol. 29, No. 1, 1975, pp. 20-24.

9. Johnson, N.L. and S. Kotz, *Continuous Distributions*, Vol. , Houghton Mifflin, Boston, 1969.

10. Kendall, M.A. and A. Stuart, *The Advanced Theory of Statistics*, Vol. 2, Hafner, 1961.

11. Miller, R.G., "The Jackknife - a Review," *Biometrika*, Vol. 61, 1974, pp. 1-15.

12. National Bureau of Standards, *Handbook of Mathematical Functions*, Washington, D.C., 1974.

13. Thoman, D.R., L.J. Bain and C.E. Antle, "Inferences on the Parameters of the Weibull Distribution," *Technometrics*, Vol. 11, No. 3, August 1969, pp. 445-460.

4. Appendix[†]

```
      SUBROUTINE ESTIMA(X,N,NUM)
C
COMMENT   THIS SUBROUTINE CONDUCTS A GRID SEARCH ON A IN INCREMENTS OF
C         DELTA
C
      INTEGER I,J,K,M,N,NUM
      REAL A,B,C,AS,BS,CS,AA(6),BB(6),CC(6),CORR(3,3),COV(3,3),D(3,3),
     2     DELTA,DEN,F,H(3),LC,LIKE(6),LLF,LOGX,MAXLLF,UC,W(15),X(N),
     3     XBAR,XSUM,Y(15)
      DATA W/.2395781703,.5601008428,.8870082629,1.22366440215,
     2        1.57444872163,1.94475197653,2.34150205664,2.77404192683,
     3        3.25564334640,3.80631171423,4.45847775384,5.27001778443,
     4        6.35956346973,8.03178763212,11.5277721009/
      DATA Y/.0933078120,.4926917403,1.2155954121,2.2699495262,
     2        3.6676227218,5.4253366274,7.5659162266,10.1202285680,
     3        13.1302824822,16.6544077083,20.7764788994,25.6238942268,
     4        31.4075191698,38.5306833065,48.0260855727/
 1      FORMAT(1H1,25X'MAXIMUM LIKELIHOOD ESTIMATION'/26X,'-------------
     2---------------'//22X,'F(X)=G(X,A,B,C)+G(X,B,A,C)+G(X,C,A,B)'//22X
     3,'G(X,T,P,R)=T*EXP(-X/T)/((T-P)*(T-R))'//4X,'N=',I5,'     SAMPLE M
     4EAN=',E13.6,'     SAMPLE VARIANCE=',E13.6//30X,'DELTA=',E13.6//12X
     5,'A=',E13.6,'     B=',E13.6,'     C=',E13.6///)
 2    FORMAT  (' SEE HYPOTHESIS',I2/////)
 3    FORMAT  (31X,'COVARIANCE MATRIX'//15X,3(E13.6,5X)//33X,2(E13.6,5X)/
     2/51X,E13.6//31X,'CORRELATION MATRIX'//15X,3(E13.6,5X)//33X,2(E13.6
     3,5X)//51X,E13.6/////)
      XSUM=0
      LOGX=0
      DO 100 I=1,N
      LOGX=LOGX+ALOG(X(I))
 100  XSUM=XSUM+X(I)
      XBAR=XSUM/N
      A=0.
      DELTA=XBAR/(3.*NUM)
      M=NUM-1
      LLF=0
      DO 150 I=1,M
      MAXLLF=LLF
      AS=A
      BS=B
      CS=C
```

[†]This set of FORTRAN subroutines computes the maximum likelihood estimates of a, b and c in f(x,a,b,c) for H_0 through H_6 in Table 1. X denotes the floating point data array, N denotes the sample size and NUM denotes the resolution DELTA = $\bar{X}$/(3*NUM) for conducting the grid search.

```
      LC=(XBAR-A)/2.
      UC=XBAR-2.*A
      CALL UPDATF(X,N,XBAR,LC,UC,A,B,C,DELTA,1,LLF)
      IF (MAXLLF.EQ.0.) MAXLLF=LLF
      IF (MAXLLF.GT.LLF) GO TO 160
 150  A=A+DELTA
 160  A=AS
      B=BS
      C=CS
C
COMMENT   ARC AND NODE SEARCH ON THE BOUNDARIES OF THE FEASIBLE REGION
C
      DO 170 I=1,3
      CALL ARC(X,N,XBAR,A,B,C,DELTA,MAXLLF,I,AA(I),BB(I),CC(I),LIKE(I))
 170  CALL NODE(N,XBAR,LOGX,A,B,C,MAXLLF,I,
     2AA(I+3),BB(I+3),CC(I+3),LIKE(I+3))
C
COMMENT   OUTPUT COMPUTATIONS FOLLOW
C
      SSQ=0
      DO 180 I=1,3
      DO 180 J=I,3
 180  D(I,J)=0
      DO 190 I=1,N
 190  SSQ=SSQ+(X(I)-XBAR)**2
      SSQ=SSQ/(N-1)
      WRITE (3,1) N,XBAR,SSQ,DELTA,A,B,C
      I=0
      IF (A.EQ.0.AND.B.LT.C) I=1
      IF (A.EQ.B.AND.B.LT.C) I=2
      IF (A.LT.B.AND.B.EQ.C) I=3
      IF (A.EQ.0.AND.B.EQ.0) I=4
      IF (A.EQ.0.AND.B.EQ.C) I=5
      IF (A.EQ.B.AND.B.EQ.C) I=6
      IF (I.EQ.0) GO TO 200
      WRITE (3,2) I
      GO TO 450
 200  DO 300 I=1,15
      CALL COMPUT(Y(I),1,B,C,A,1,H(1),F)
      CALL COMPUT(Y(I),1,A,C,B,1,H(2),F)
      CALL COMPUT(Y(I),1,A,B,C,1,H(3),F)
      F=EXP(F)
      DO 300 J=1,3
      DO 300 K=J,3
 300  D(J,K)=D(J,K)+H(J)*H(K)*F*W(I)
      DEN=(D(1,1)*D(2,2)*D(3,3)+2.*D(1,2)*D(2,3)*D(1,3)
     2-D(2,2)*D(1,3)**2-D(3,3)*D(1,2)**2-D(1,1)*D(2,3)**2)*N
```

```
      COV(1,1)=(D(2,2)*D(3,3)-D(2,3)**2)/DEN
      COV(1,2)=(-D(1,2)*D(3,3)+D(1,3)*D(2,3))/DEN
      COV(1,3)=(D(1,2)*D(2,3)-D(1,3)*D(2,2))/DEN
      COV(2,2)=(D(1,1)*D(3,3)-D(1,3)**2)/DEN
      COV(2,3)=(-D(1,1)*D(2,3)+D(1,2)*D(1,3))/DEN
      COV(3,3)=(D(1,1)*D(2,2)-D(1,2)**2)/DEN
      DO 400 I=1,3
      DO 400 J=I,3
 400  CORR(I,J)=COV(I,J)/SQRT(COV(I,I)*COV(J,J))
      WRITE (3,3)     COV(1,1),COV(1,2),COV(1,3),COV(2,2),COV(2,3),
     2COV(3,3),CORR(1,1),CORR(1,2),CORR(1,3),CORR(2,2),CORR(2,3),
     3CORR(3,3)
C
COMMENT CHECK HYPOTHESES 1,2 AND 3
C
      CALL HYP123(X,N,XBAR,LOGX,DELTA,MAXLLF,AA,BB,CC,LIKE)
C
COMMENT   CHECK HYPOTHESES 4,5 AND 6
C
      DO 500 I=1,3
 500  CALL ERLANG(N,XBAR,LOGX,MAXLLF,I)
      END



      DOUBLE PRECISION FUNCTION G(Y,THETA,PHI,RHO)
C
COMMENT   COMPUTES THETA*EXP(-Y/THETA)/((THETA-PHI)*(THETA-RHO))
C
      REAL PHI,RHO,S,THETA,Y
      REAL*8 ARG,CHECK,Z,ZZ,ZZZ
      G=0
      IF (THETA.EQ.0.) RETURN
      ARG=Y/THETA
      S=1.
      Z=THETA/((THETA-PHI)*(THETA-RHO))
      IF (Z.LT.0.) S=-S
      IF (ARG.LE.174.673) GO TO 25
 10   ARG=-ARG+DLOG(DABS(Z))
      IF (ARG.LT.-180.218) RETURN
      G=S*DEXP(ARG)
      RETURN
 25   ZZZ=DEXP(ARG)
      CHECK=(10D-78)*ZZZ
      ZZ=DABS(Z)
      IF (ZZ.LT.CHECK) RETURN
      G=Z/ZZZ
      RETURN
      END
```

```
      SUBROUTINE UPDATE(X,N,XBAR,LC,UC,A,B,C,DELTA,J,LLF)
C
COMMENT   PERFORMS BINARY SEARCH ON C FOR GIVEN A
C
      INTEGER J,N
      REAL A,B,C,CC,DEL,DELTA,DERIVC,LC,LLF,UC,W(12),X(N),XBAR,V(16)
      DATA V/0.,0.,0.,1.,1.,0.,0.,0.,0.,0.,1.,0.,0.,0.,0.,-2./
      DATA W/1.,1.,.5,0.,-1.,0.,0.,0.,-1.,-1.,-.5,1./
      CC=(LC+UC)/2.
 100  C=CC
      B=XBAR*W(J)+A*W(J+4)+C*W(J+8)
      A=XBAR*V(J)+A*V(J+4)+B*V(J+8)+C*V(J+12)
      CALL COMPUT(X,N,A,B,C,J,DERIVC,LLF)
      IF (DERIVC.GE.0) LC=C
      IF (DERIVC.LE.0) UC=C
      CC=(LC+UC)/2.
      DEL=ABS(C-CC)
      IF (DEL.GT.DELTA) GO TO 100
      RETURN
      END


      SUBROUTINE COMPUT(Y,M,T,P,R,J,DERIVC,LLF)
C
COMMENT   COMPUTES LOGLIKELIHOOD DERIVATIVE WITH RESPECT TO C
C
      INTEGER I,J,K,M
      REAL DERIVC,LLF,P,R,T,Y(M)
      REAL*8 F,G,GP,GR,GT
      LLF=0
      DERIVC=0
      GO TO (100,200,300,400), J
 100  DO 150 K=1,M
      GT=G(Y(K),T,P,R)
      GP=G(Y(K),P,T,R)
      GR=G(Y(K),R,T,P)
      F=GT+GP+GR
      LLF=LLF+DLOG(F)
      GT=GT/F
      GP=GP/F
      GR=GR/F
 150  DERIVC=DERIVC+GR*(1./R+Y(K)/R**2+1./(P-R)+1./(T-R))
     2+GP/(P-R)+GT/(T-R)
      RETURN
 200  DO 250 K=1,M
      GP=G(Y(K),P,R,0)
```

```
      GR=G(Y(K),R,P,0)
      F=GP+GR
      LLF=LLF+DLOG(F)
      GP=GP/F
      GR=GR/F
 250  DERIVC=DERIVC+GP/(P-R)+GR*(Y(K)/R**2-1./(R-P))
      RETURN
 300  DO 350 K=1,M
      GP=G(Y(K),P,R,R)
      GR=G(Y(K),R,P,P)
      F=GR+GP*((P-R)*Y(K)-P*R)/P**2
      LLF=LLF+DLOG(F)
      GP=GP/F
      GR=GR/F
 350  DERIVC=DERIVC+GR*(1./R+Y(K)/R**2-2./(R-P))
     2+GP*((R-P)*Y(K)+P*(P+R))/((R-P)*P**2)
      RETURN
 400  DO 450 K=1,M
      GT=G(Y(K),T,R,R)
      GR=G(Y(K),R,T,T)
      F=GT+GR*((R-T)*Y(K)-T*R)/R**2
      LLF=LLF+DLOG(F)
      GT=GT/F
      GR=GR/F
 450  DERIVC=DERIVC+2.*GT/(T-R)
     2+GR*(((R-T)*Y(K)-T*R)*(Y(K)/R**2-1./R-2./(R-T))+Y(K)-T)/R**2
      RETURN
      END
```

```
      SUBROUTINE ARC(X,N,XBAR,A,B,C,DELTA,MAXLLF,I,AA,BB,CC,LIKF)
C
COMMENT COMPUTES SOLUTIONS FOR ARCS AND APPLIES TO HYPOTHESES 1,2 AND 3
C
      INTEGER I,N
      REAL A,AA,B,BB,C,CC,DELTA,LIKF,MAXLLF,U(6),X(N),XBAR
      DATA U/.5,.333333,.333333,1.,1.,.5/
      CALL UPDATE(X,N,XBAR,XBAR*U(I),XBAR*U(I+3),AA,BB,CC,DELTA,I+1,
     2LIKF)
      IF (MAXLLF.GT.LIKF) RETURN
      A=AA
      B=BB
      C=CC
      MAXLLF=LIKF
      RETURN
      END
```

```
      SUBROUTINE NODE(N,XBAR,LOGX,A,B,C,MAXLLF,I,AA,BB,CC,LIKE)
C
COMMENT   COMPUTES SOLUTIONS FOR NODES AND APPLIES TO HYPOTHESES 4,5,6
C
      INTEGER I,N
      REAL A,AA,B,BB,C,CC,LIKE,LOG2,LOGX,MAXLLF,XBAR,W(9)
      DATA W/0.,0.,.333333,0.,.5,.333333,1.,.5,.333333/,LOG2/.693147/
      LIKE=-N*(I*(1.+ALOG(XBAR/I))+3.*W(I)*LOG2)+(I-1)*LOGX
      AA=XBAR*W(I)
      BB=XBAR*W(I+3)
      CC=XBAR*W(I+6)
      IF (LLF.GT.LIKE) RETURN
      A=AA
      B=BB
      C=CC
      MAXLLF=LIKE
      RETURN
      END



      SUBROUTINE HYP123(X,N,XBAR,LOGX,DELTA,MAXLLF,AA,BB,CC,LIKE)
C
COMMENT   PERFORMS OUTPUT ANALYSES FOR HYPOTHESES 1,2 AND 3
C
      INTEGER I,J,K(9),KA,KB,KC,L,N
      REAL A,AA(6),B,BB(6),C,CC(6),CBB,CCC,CBC,CCBB,CCCC,CCBC,DELTA,
     2     D(2,2),DFN,F,HB,HC,LIKE(6),LOGX,LRATIO,MAXLLF,U(6),W(15),
     3     X(N),XBAR,Y(15)
      DATA K/1,1,2,2,3,3,1,2,1/
      DATA U/.5,.333333,.333333,1.,1.,.5/
      DATA W/.2395781703,.5601008428,.8870082629,1.2236640215,
     2       1.5744872163,1.9475197653,2.3415020566 4,2.7740019268 3,
     3       3.2556413464 0,3.8063117142 3,4.4584777538 4,5.2700177844 3,
     4       6.3595634697 3,8.0317876321 2,11.527721009/
      DATA Y/.0933078120,.4926917403,1.2155954121,2.2639495262,
     2       3.6676227218,5.4253366274,7.5659162266,10.1202285680,
     3       13.1302824822,16.6544077083,20.7764788994,25.6238942268,
     4       31.4075191698,38.5106933065,48.0260855727/
 1    FORMAT(' HYPOTHESIS 1:   A=0, B<=C'//)
 2    FORMAT(' HYPOTHESIS 2:   A=B<=C'//)
 3    FORMAT('1HYPOTHESIS 3:   A<=B=C'//)
 4    FORMAT(15X,'A=',E13.6,5X,'B=',E13.6,5X,'C=',E13.6//'       SEE HYP
     2OTHESIS ',I2/////)
```

```
5     FORMAT(22X,'B=',E13.6,5X,'C=',E13.6//9X,'VAR(B)=',F13.6,5X,'VAR(C)
     2=',E13.6,5X,'COV(B,C)=',E13.6//28X,'CORR(B,C)=',E13.6//25X,'LIKELI
     3HOOD RATIO=',E13.6/////)
6     FORMAT(22X,'A=',E13.6,5X,'C=',F13.6//9X,'VAR(A)=',E13.6,5X,'VAR(C)
     2=',E13.6,5X,'COV(A,C)=',E13.6//28X,'CORR(A,C)=',E13.6//25X,'LIKELI
     3HOOD RATIO=',E13.6/////)
      DO 500 I=1,3
      IF (I.EQ.1) WRITE (3,1)
      IF (I.EQ.2) WRITE (3,2)
      IF (I.EQ.3) WRITE (3,3)
      L=0
      KA=K(I)
      KB=K(I+3)
      KC=K(I+6)
      L=I
      DO 100 J=KA,KB,KC
      JJ=J+3
      IF (LIKE(I).LT.LIKE(JJ)) L=JJ
      A=AA(L)
      B=BB(L)
      C=CC(L)
100   LIKE(I)=LIKE(L)
      L=0
      IF (A.EQ.0.AND.B.EQ.0) L=4
      IF (A.EQ.0.AND.B.EQ.C) L=5
      IF (A.EQ.B.AND.B.EQ.C) L=6
      IF (L.LT.4) GO TO 150
      WRITE (3,4) A,B,C,L
      GO TO 500
150   LRATIO=EXP(LIKE(I)-MAXLLP)
      DO 175 J=1,2
      DO 175 L=J,2
175   D(J,L)=0
      DO 475 J=1,15
      GO TO (200,300,400), I
200   CALL COMPUT(Y(J),1,A,C,B,2,HB,F)
      CALL COMPUT(Y(J),1,A,B,C,2,HC,F)
      GO TO 450
300   CALL COMPUT(Y(J),1,C,B,A,4,HB,F)
      CALL COMPUT(Y(J),1,A,B,C,3,HC,F)
      GO TO 450
400   CALL COMPUT(Y(J),1,C,B,A,3,HB,F)
      CALL COMPUT(Y(J),1,A,B,C,4,HC,F)
450   F=EXP(F)
      D(1,1)=D(1,1)+HB**2*F*W(J)
      D(2,2)=D(2,2)+HC**2*F*W(J)
475   D(1,2)=D(1,2)+HB*HC*F*W(J)
      DEN=(D(1,1)*D(2,2)-D(1,2)**2)*N
```

```
      CBB=D(2,2)/DEN
      CCC=D(1,1)/DEN
      CBC=-D(1,2)/DEN
      CCBC=-D(1,2)/SQRT(D(1,1)*D(2,2))
      IF (I.EQ.1) WRITE (3,5) B,C,CBB,CCC,CBC,CCBC,LRATIO
      IF (I.EQ.2) WRITE (3,5) B,C,CBB,CCC,CBC,CCBC,LRATIO
      IF (I.EQ.3) WRITE (3,6) A,C,CBB,CCC,CBC,CCBC,LRATIO
 500  CONTINUE
      RETURN
      END


      SUBROUTINE ERLANG(N,XBAR,LOGX,MAXLLF,I)
C
COMMENT   PERFORMS OUTPUT ANALYSES FOR HYPOTHESES 4,5 AND 6
C
      INTEGER I,N
      REAL C,CHISQ,DF,LC,LLF,LOG2,LOGX,LRATIO,MAXLLF,UC,W(3),XBAR
      DATA W/0.,0.,1./, LOG2/.693147/
 1    FORMAT(' HYPOTHESIS 4:   A=B=0'//)
 2    FORMAT(' HYPOTHESIS 5:   A=0, B=C'//)
 3    FORMAT(' HYPOTHESIS 6:   A=B=C'//)
 4    FORMAT(32X,'C=',E13.6//8X,'.95 LOWER POINT=',E13.6,5X,'.95 UPPER P
     2OINT=',E13.6//25X,'LIKELIHOOD RATIO=',E13.6/////)
      C=XBAR/I
      DF=2.*I*N
      LC=DF*C/CHISQ(DF,.975)
      UC=DF*C/CHISQ(DF,.025)
      LLF=-N*(I*(1.+ALOG(C))+W(I)*LOG2)+(I-1)*LOGX
      LRATIO=EXP(LLF-MAXLLF)
      IF (I.EQ.1) WRITE (3,1)
      IF (I.EQ.2) WRITE (3,2)
      IF (I.EQ.3) WRITE (3,3)
      WRITE (3,4) C,LC,UC,LRATIO
      RETURN
      END
```

```
      FUNCTION CHISQ(DF,P)
COMMENT   COMPUTES CRITICAL VALUE OF CHI-SQUARE FOR PROBABILITY P
C         AND DF DEGREES OF FREEDOM
      INTEGER I
      REAL C(3),D(3)
      REAL DF,P,Q,T,XP,NUM,DEN,Y,SQDF,SQHALF,Y2,Y3,Y4,Y5,Y6,Y7,H(7)
      DATA C/2.515517,.802853,.010328/,D/1.432788,.189269,.001308/
      NUM=0
      DEN=1.
      Q=P
      IF (P.LE..5) GO TO 5
      Q=1.-P
 5    T=SQRT(ALOG(1./Q**2))
      DO 10 I=1,3
      NUM=NUM+C(I)*T**(I-1)
 10   DEN=DEN+D(I)*T**I
      XP=T-NUM/DEN
      IF (P.GE..5) GO TO 15
      XP=-XP
 15   Y=XP
      SQDF=SQRT(DF)
      SQHALF=SQRT(.5)
      Y2=Y*Y
      Y3=Y*Y2
      Y4=Y3*Y
      Y5=Y4*Y
      Y6=Y5*Y
      Y7=Y6*Y
      H(1)=Y/SQHALF
      H(2)=2.*(Y2-1.)/3.
      H(3)=(Y3-7.*Y)*SQHALF/9.
      H(4)=-(6*Y4+14.*Y2-32.)/405.
      H(5)=(9.*Y5+256.*Y3-433.*Y)*SQHALF/4860.
      H(6)=(12.*Y6-243.*Y4-923.*Y2+1472.)/25515.
      H(7)=-(3753.*Y7+4353.*Y5-289517.*Y3-289717.*Y)*SQHALF/9185400.
      CHISO=1.
      DO 20 I=1,7
 20   CHISQ=CHISO+H(I)/SQDF**I
      CHISQ=CHISQ*DF
      RETURN
      END
```

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: TR-76-7	2. GOVT ACCESSION NO.
	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle): Maximum Likelihood Estimation of the Distribution of the Sum of Three Independent Exponential Random Variables.	5. TYPE OF REPORT & PERIOD COVERED: Technical Report
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): George S. Fishman	8. CONTRACT OR GRANT NUMBER(s): N00014-67-A-0321-0008
9. PERFORMING ORGANIZATION NAME AND ADDRESS: University of North Carolina, Chapel Hill, North Carolina 27514	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS: Operations Research Program, Office of Naval Research, Arlington, Virginia 22217	12. REPORT DATE: May 1976
	13. NUMBER OF PAGES: 25
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS. (of this report)
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Distribution of this document is 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)

Binary Search
Exponential Random Variables
Grid Search
Likelihood Ratio
Maximum Likelihood Estimation
Regenerative Processes
Simulation

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

This paper describes a procedure for computing the maximum likelihood estimates of the parameters of the distribution of the sum of three independent exponential random variables. By fitting sample interevent time data from a real system to this distribution, one can create a simulation of the system that exploits the regenerative representation of queueing systems to analyze the simulation's output by relatively elementary statistical methods. The paper also describes computation of the sample asymptotic covariance matrix and an →

259500

20. Abstract cont.

implementation of the likelihood ratio for testing six hypotheses that are special cases of interest. A set of FORTRAN subroutines for executing these procedures appears in the Appendix.