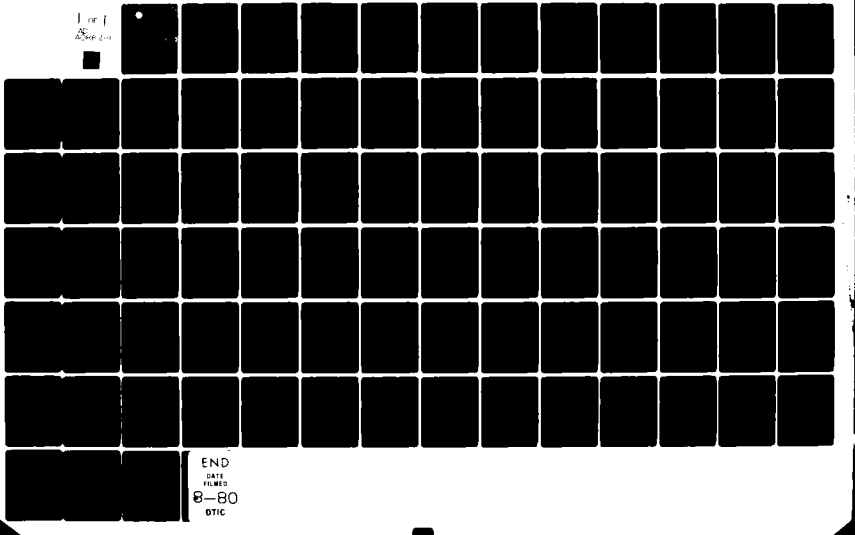
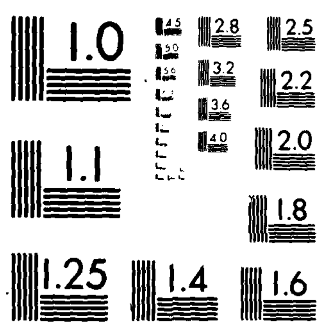
AD-A086 281 SOUTHERN METHODIST UNIV DALLAS TEX DEPT OF ELECTRICA--ETC F/G 9/3
ESTIMATION OF CONFIDENCE LIMITS FOR TESTING LARGE LOGIC NETWORK--ETC(U)
MAY 80 J L FIKE, C H KAPADIA, B PEIKARI AFOSR-77-3461
UNCLASSIFIED AFOSR-TR-80-0504 NL



END
DATE FILMED
8-80
DTIC



MICROCOPY RESOLUTION TEST CHART

LEVEL II

ESTIMATION OF CONFIDENCE LIMITS FOR TESTING LARGE LOGIC NETWORKS

FINAL REPORT

Prepared Under AFOSR Grant No. 77-3461

Prepared by

John Fike and
C. H. Kapadia
Co-Principal Investigators

May 1980

SOUTHERN METHODIST UNIVERSITY
DALLAS, TEXAS 75275

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

(19) REPORT DOCUMENTATION PAGE	READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER: (18) AFOSR-TR-80-0504	2. GOVT ACCESSION NO. AD-A086 281
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle): (6) ESTIMATION OF CONFIDENCE LIMITS FOR TESTING LARGE LOGIC NETWORKS	5. TYPE OF REPORT & PERIOD COVERED: (9) Final
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): (10) John L. Fike, K. Kavipurapu, C. H. Kapadia, B. Peikari	8. CONTRACT OR GRANT NUMBER(s): (15) AFOSR 77-3461
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Southern Methodist University, Dept. of Electrical Engineering, Dallas, TX 75275	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: (16) 61102F (17) 2304/A6
11. CONTROLLING OFFICE NAME AND ADDRESS: Air Force Office of Scientific Research/NM, Bolling AFB, Washington, DC 20332	12. REPORT DATE: (11) May 1980
	13. NUMBER OF PAGES: 83
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): (12) 83	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)

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

This report describes a methodology and means for accurately determining confidence limits for the reliability of large digital networks, without exhaustively exercising all possible input sequences or simulating all logic faults in the network. The report is divided into two parts. In the first part, some mathematical models to estimate the reliability of digital circuits are presented. A heuristic method of assigning weights to faults depending on their "importance" in a given circuit is described. The models presented can be used to: (a) predict the reliability of a circuit, (b) evaluate test sequences and (c) develop

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20. Abstract cont.

more accurate reliability models of (redundant) fault tolerant computers. The second part of the report deals with the statistical methods of estimating confidence limits.

Estimation of Confidence Limits For Testing Large Logic Networks

Final Report

Prepared Under AFOSR Grant No. 77-3461

for

Department of the Air Force

Air Force Office of Scientific Research

Bolling AFB, Washington D. C.

Prepared by

J. L. Fike

C. H. Kapadia

B. Peikari

K. Kavipurapu

Southern Methodist University

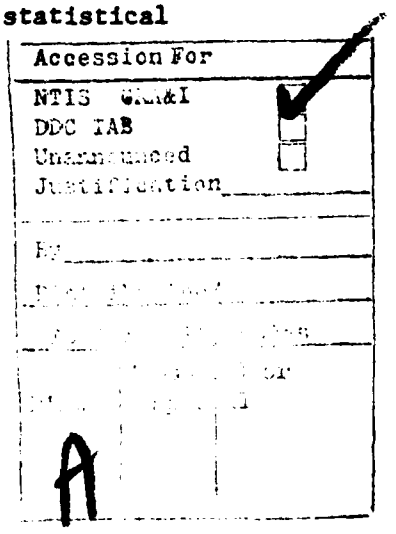
Dallas, Texas 75275

May 1980

Table of Contents

ABSTRACT

This report describes a methodology and means for accurately determining confidence limits for the reliability of large digital logic networks, without exhaustively exercising all possible input sequences or simulating all logic faults in the network. The report is divided into two parts. In the first part, some mathematical models to estimate the reliability of digital circuits are presented. A heuristic method of assigning weights to faults depending on their "importance" in a given circuit is described. The models presented can be used to: (a) predict the reliability of a circuit, (b) evaluate test sequences and (c) develop more accurate reliability models of (redundant) fault tolerant computers. The second part of the report deals with the statistical methods of estimating confidence limits.



Chapter I

INTRODUCTION

With the increased complexity of current digital systems, reliability considerations have become increasingly important. Being physical devices, digital circuits are subject to failure. Although current technologies employed to construct digital systems are more reliable than earlier technologies, to a great extent the resulting decrease in the failure rate of individual components has been offset by the increased complexity of today's circuits. This is one of the main reasons for the increased interest, both in industry and academia, in the subjects of maintenance and reliability of digital systems.

A fault or failure of a digital network can be defined as a physical defect of one or more components in the network which causes it to behave differently from the original systems (Su [74]). In digital systems, typical maintenance goals deal with the rapid detection, location and repair of any system faults. In many digital systems involving real-time processes, such as telephone switching networks and aircraft or spacecraft flight controls, it is desirable to continuously monitor, exercise and test the system in order to determine whether the system is performing as desired. Such monitoring may enable automatic detection of failures via periodic testing or through the use of codes and checking circuits (e.g. self-testing and self-checking circuits) or may enable continuous operation under failure (i.e., fault tolerence) and automatic repair via switching networks (e.g. stand-by spares).

One way to determine whether a fault exists in a circuit is to exercise the network against all possible input sequences and compare the results with expected output. A second method is to insert faults, physical or modelled,

into a copy of the network, then generate tests that detect the presence of these faults. For a complex digital system such as an airborne computer, it is not only impossible to apply all possible input sequences but also impractical to model all possible faults. A more practical and accepted approach is to generate input patterns that detect a certain percent of faults. The generation and evaluation of such test sequences have been the subjects of much investigation (see for example Friedman [71] or Breuer [76]).

As the average number of IC's on a board increases, the difficulty of adequately testing the board increases, perhaps exponentially. It can be shown that with normal rejection rates of incoming IC's, boards of 50 to 100 IC's will always contain at least one bad IC prior to testing (Fike [72]).

As things stand now, checking out a $10 microprocessor chip may require an investment by the user of upwards of 1000 times that amount in test gear (e.g. Fairchild's Sentry VII) - and he still may not know if the chip will do the job it is supposed to do (Vodovoz [75]). Both user and manufacturer come face to face with the problems of seeing if the chips work and at this moment, no one is fully satisfied with present chip check-out techniques. The hardest hit, though, is the end user who measures his chip needs in hundreds, and who can not afford the sophisticated test equipment that can give him better answers than those he gets with his own home brewed test methods.

Two truisms must be recognized at this point:

1. Testing does not add to product quality, it merely evaluates the quality already present.
2. The test function normally requires a larger expenditure for equipment than needed for any other part of the production organization.

Another way of expressing the first statement is to say that the quality must be built in through use of good components and good workmanship. The second statement indicates that the question of the amount of testing to be performed must be very carefully considered and answered. The amount of testing which should be performed is that minimum which clearly demonstrates that the product performs or fails to perform as specified. This rather vague definition must be expanded greatly so that all responsible personnel have a reasonably accurate concept as to what constitutes a 'proper minimum test' and also what constitutes the economics of testing. One such attempt appeared in Watkins [70].

Most users demand that the test equipment detect troubles which are not present at the time of testing but which may occur at some future time. To achieve such tight tolerences, one has to test his system on a regular basis. The cost of such testing can be prohibitively expensive. So, we need a procedure which enables the end user to evaluate the trade-offs between the frequency of testing (consequently, the cost of testing) and the "confidence" he can have in his system. This is the subject of our investigations.

This report describes a methodology and means for accurately determining confidence limits for the reliability of large digital logic networks, without exhaustively exercising all possible input sequences or simulating all logic faults in the network. This report is divided into two parts. In the first part, some mathematical models to estimate the reliability of digital circuits are presented. A heuristic method of assigning weights to faults depending on their "importance" in a given circuit is described. The models presented can be used to

1. predict the reliability of a circuit

2. evaluate test sequences

3. develop more accurate reliability models of (redundant) fault tolerent computers.

The second part deals with the statistical methods of estimating confidence limits.

Chapter II

TEST SEQUENCES

Test generation is the process of finding the set of input patterns for a digital circuit which will either verify that the circuit is operating correctly or else provide some information on the nature of the failure. The set of such input patterns is often designated as a test sequence (or simply, a test). It is desirable that the test

1. be reasonable in size
2. be produced at reasonable expense, and yet,
3. detect a maximum number of faults.

Many algorithms have been proposed for generating tests for digital circuits. Some of the procedures make use of an algebraic description of the circuit under consideration. Others directly utilize the gate level circuit topology and functional description. In this section we will survey briefly some of these algorithms. In almost all these studies, a common assumption is that a logic element can only fail by sticking at zero or by sticking at one.

The test generation procedures that have been evolved so far can be broadly classified as either deterministic or probabilistic. Examples of deterministic techniques include the Boolean difference method (Sellars [68]), the one dimensional path sensitization method (Armstrong [66]) and D-algorithm (Roth [66]). All of these methods offer high fault coverage (i.e., each test pattern detects a loarge number of faults) and reasonable size test sequences. But, they are computationally complex and limited to special classes of circuits. For example, the D-algorithm uses cubical algebra and

requires extensive programming to automate the procedure. Moreover, the D-algorithm is originally designed for combinational circuits. Extensions to these methods to handle sequential circuits have evolved in recent years, but the complexity prohibits their use.

In the random technique, a candidate test is chosen using a random number generator. If the candidate test detects new faults not detected by previous tests, it is added to the test set; otherwise it is discarded. The figure of merit of a candidate test can be defined as the ratio of the number of new faults detected to the number of faults to be detected. A candidate test may be discarded if its figure of merit is less than a specified threshold value.

There are two major categories of Automatic Test Equipments (ATE), called stored program ATE and comparison ATE. Stored program testers usually contain a mini computer and back-up storage as disk and test sequences stored vector by vector or as a high level program interpreted by the computer. The stored program ATE typically also stores the responses and a fault dictionary, which are usually generated by simulation. The actual test sequences can be obtained using either a deterministic or random procedure.

Comparison ATE employs pseudo-random patterns as test vectors. Here two circuits, the Unit Under Test (UUT) and a known good copy of the circuit (denoted by C^*) are inserted into the ATE (Figure 1).

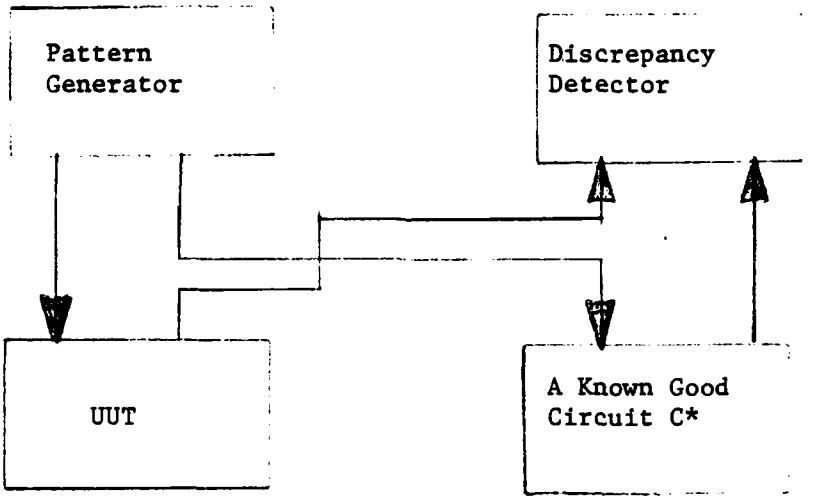


Figure 1. Comparison ATE

The pattern generator applies several million pseudo random patterns to both the UUT and C^*. The outputs of the circuits are then compared by the discrepancy detector. A mismatch indicates a fault in UUT. The advantages of comparison type ATE are 1. very fast generation of test patterns, 2. absence of an expensive computer and storage. However, a known good copy of the circuit is needed.

Although random test patterns can be generated very inexpensively, this technique becomes progressively inefficient when attempting to detect more deeply embedded faults (Parker [75a]). Adaptive (Parker [75a]) and Weighted (Schnurmann [75]) random test generations have been developed to improve the efficiency of random test generation.

For a complex digital network, one is usually satisfied with tests that detect a certain percent of the modelled faults (say 95%). As this number approaches 100%, the number of input patterns required to detect these faults approaches a very large number. So, most end users are satisfied with tests that check a percentage of faults.

Let G denote the set of all modelled faults and let Q be the set of faults an user want to detect ($Q \subseteq G$). Typically, if an end user is satisfied with tests that detect 95% of the modelled faults, then 95% of G constitutes Q. A test sequence T is then generated to detect the faults in Q and the circuit is exercised against all the input patterns in T periodically to assure reliable operation of the circuit.

Let T1, T2,...,Tn be a number of test sequences detecting faults in Q1, Q2,...,Qn where each Qi is 95% (or any given percent) of G. The actual set of faults detected by each of these tests Ti can be determined before hand by probing. By using a fault dictionary, a fault may be identified within its equivalent class. The set of physical faults associated with this equivalent class may be distributed over many components, such as IC's. In order to determine the exact location of a physical fault, probing techniques are required. It should be clear that by interchanging these test Ti periodically, say T1 during the first testing, T2 when the circuit is tested second time and so on, the reliability associated with the circuit can be enhanced appreciably (Figure 2).

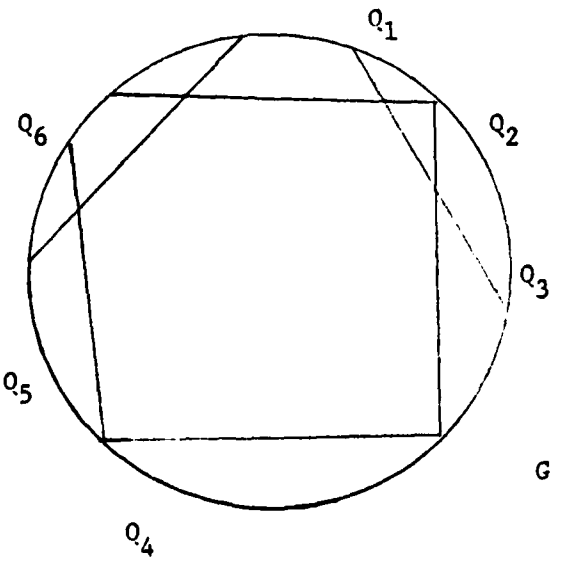


Figure 2. Sets of Faults

PROBABILITY OF DETECTING A FAULT

Parker [75b] has defined signal probability as follows:

Definition: The probability of a signal denoted as

$$a = P(A=1) \tag{1}$$

for a signal A is the probability that signal A equals 1. Similarly, the probability that the signal equals 0 is given by

$$P(A=0) = 1 - P(A=1) = 1 - a \tag{2}$$

Boolean operations NOT, AND, OR can be applied on probabilities.

(a) Boolean Negation (NOT) corresponds to the probability expression

$$b = 1 - a \tag{3}$$

(b) Boolean AND of two independent signals A and B in the expression C = A.B corresponds to the probability expression

$$c = a \cdot b \tag{4}$$

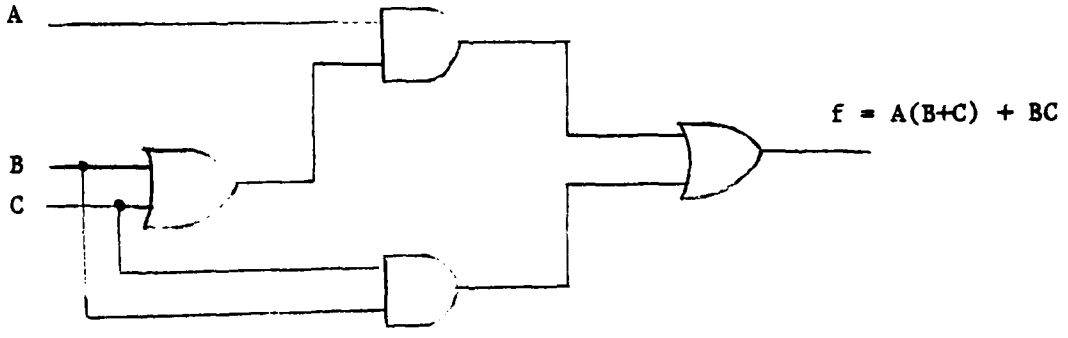
(c) Boolean OR of two independent signals A and B in the expression C = A + B corresponds to the probability expression

$$c = a + b \tag{5}$$

Let $f(x1,x2,\ldots,xn)$ denote the function realized by a combinational circuit. A logical fault α changes the function realized to f_α $(x1,x2,\ldots xn)$. Using the theory of boolean difference method (Sellers [68]), the fault α is detected by an input for which $f + f_\alpha$ is 1. Using the probability expression for NOT, AND, OR operations, the probability $P(f + f_\alpha = 1)$ can be evaluated as a function of the input signal probabilities. This is the probability that fault α will be detected. It is often convenient to assign

equal probabilities to input signals. This is known as bundling (Parker [75c]).

To see how the probability of detecting a fault can be used in random testing let us consider the circuit in Figure 3.



(6)

$$f = ABC + \bar{A}BC + A\bar{B}C + AB\bar{C} \quad (6)$$

Figure 3 An Example - Probability of detecting a fault.

Let α be the fault that A is stuck at 1. Then

$$f_\alpha (A,B,C) = B + C \quad (7)$$

$$f + f_\alpha = B\bar{C} + \bar{A}BC \quad (8)$$

$$P(f + f_\alpha = 1) = b + c - ac - 2bc + abc \quad (9)$$

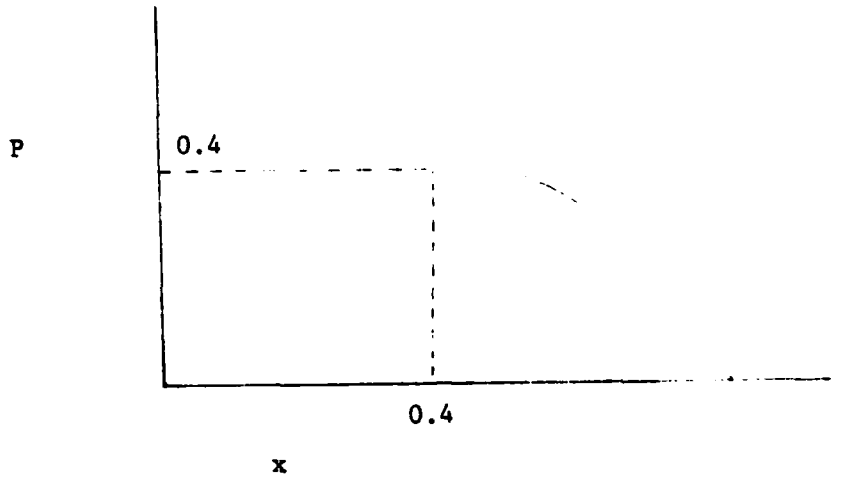
Assuming equal probabilities to a, b, and c, i.e.,

$a = b = c = x$

the probability of detecting fault α is given by

$$P(f + f_\alpha = 1) = 2x - 3x^2 + x^3 \quad (10)$$

Since x is a probability, the above function is plotted for values of x between 0 and 1.



This function $P(f + f_\alpha = 1)$ in terms of the input signal probabilities can be used to evaluate the values of the input signal probabilities for any desired probability of detecting the fault α. Theses values can be used to control the random input patterns generated leading to adoptive test generation. For example, from the above plot, the input signal probabilities (i.e., x) must be set to 0.4 in order to achieve a probability of detecting the fault of 0.4.

DEPENDENCY OF RANDOM TEST GENERATION ON THE LOGIC DEPTH

The Monte Carlo method was used extensively for testing the circuits of ILLIAC IV (Moreno [72], Agrawal [72]). Some of the interesting results obtained during the testing are given here. It was found that a circuit with 437 lines was completely tested (all s.a.0 and s.a.1 faults with single fault assumption) with 56 random input patterns. Another circuit with 89 lines required 210 patterns for complete testing while for a third circuit with 115 lines, the test generation was not complete even after 2000 patterns. These results lead to the conclusion that the computation time does not depend only on the number of lines or faults. Agrawal [75a] showed

that the number of random inputs required for a complete test of a circuit depends on the logic depth, i.e., the number of levels in the circuit. In Agrawal [75b], an expression for the probability of sensitizing a path upto the primary output through L levels, P(L) is derived. Then, the probability of sensitizing a path through L levels by at least one out of M independent patterns P(L,M) is given by

$$P(L,M) = 1 - [1 - P(L)]^M \tag{11}$$

Solving the above equation for M,

$$M = \frac{\log[1 - P(L,M)]}{\log[1 - P(L)]} \tag{12}$$

This equation can be used to estimate the number of random input patterns required to detect all faults in a circuit.

Experimental results have shown that equation (12) gives a good estimate of the number of random input patterns required for a complete check of combinational circuits (Agrawal [75b]). Although these results are derived for simple tree structures consisting of n input NAND gates, the model can be applied to any combinational circuit since an equivalent NAND trees can be constructed for the given circuit very easily.

CONFIDENCE LEVEL

In Agrawal [75b], the term confidence is used to mean the probability of sensitizing a path through L levels by at least one of the M random input patterns in a test set. Another way of describing this meaning is to say that the confidence of a test sequence is the probability of detecting a fault since a fault is detected if a path from the site of the fault to a primary output can be sensitized.

In another attempt to define confidence level, Shedletsky [77] derived an expression for latency intervals in a circuit. Error latency can be defined as the delay between the occurrence of a fault and the first error in the output. Latency interval of a circuit is the maximum of the minimum number of input patterns necessary to achieve a given probability of detecting a fault. Shedletsky notes that the required length of a random test (i.e., the number of input sequences) to achieve a given confidence level is equal to the latency interval of the circuit. A more detailed discussion of error latency is included in a later section. Thus, in this definition, confidence level is related to the number of test patterns in that test set.

In both of the above attempts, confidence level is defined in terms of a test sequences. However, a more useful definition of confidence level should relate to the reliability of the circuit itself. A meaningful definition should include mean time between failures (mtbf) of the circuit due to logical and physical faults. Such a definition can be used to schedule periodic check-outs in order to achieve a desired confidence in the circuit. In the next section two definitions of reliability are given. The authors believe that these definitions are more useful.

RELIABILITY ASSOCIATED WITH A TEST SEQUENCE

Here two definitions of reliability are presented. The first definition accounts for the faults that are detected by the test sequence under consideration. This definition is then extended to include modelled faults that are not detected by the test sequence and even faults that can not be modelled.

Definition 1: The reliability R_T of a circuit associated with a test sequence T which detects the set of faults Q = (g1,g2,...,gn) is a function

of time and is given by

$$R_T = \prod_{i=1}^{n} [(1-R_i \int_{to}^{t} fi \cdot dt)] \qquad (13)$$

assuming that the circuit has passed the test T at time t0 and where

fi is the probability density function of the fault gi with a mean of $1/\lambda_i$

and ki is a constant which describes the dependence of reliability R_T on the fault gi, $0 < ki < 1$.

The term $(1 - \int_{to}^{t} fi \cdot dt)$ in the above equation gives the probability that the fault gi does not exist at time t since $\int_{to}^{t} fi \cdot dt$ is the cumulative probability that fault gi is present at time t assuming that gi did not exist at time t0.

Typically, all ki's are set equal to 1, however, an user can bias the reliability in favor of some of the faults depending on their importance in his circuit. These constants are in line with the weighted random test pattern generation (Schnurmann [75]) where a weight to each signal is assigned according to its importance in the circuit.

Equation (13) describes some kind of decay function for R_T and Figure 5 shows the general shape of such a curve. By applying the test T periodically, the reliability R_T can be restored to the maximum (provided the circuit passes the test T each time). In such cases, the curve has a sawtooth form (Figure 6). The frequency at which the circuit must be tested depends on the minimum reliability R_T.

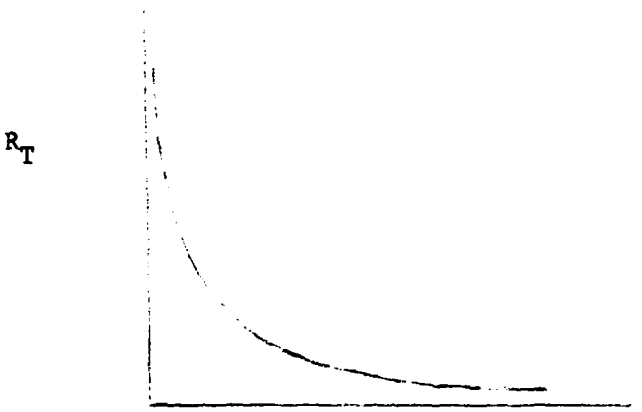


Figure 5 Reliability Associated with a Test.

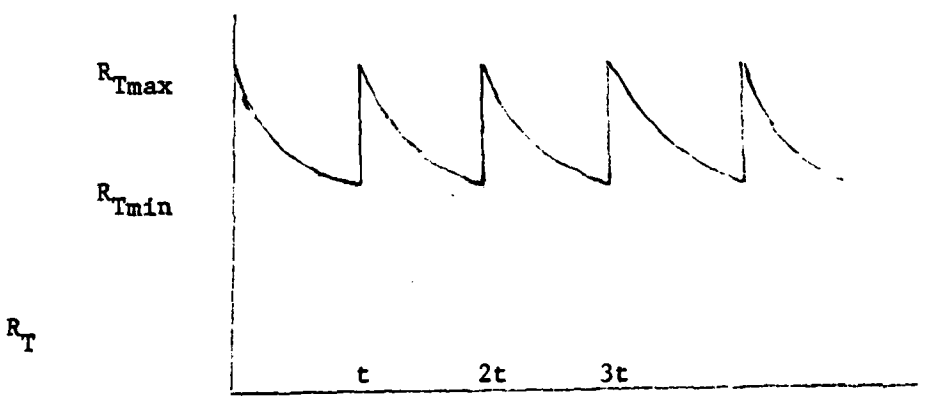


Figure 6 Periodic Testing Restores Reliability to its Maximum

The above definition takes into account only the faults that are detected by the test T. But, intuitively, it should be obvious that the reliability of the circuit should be lower than that predicted by the above equation due to the faults that are not detected by test T. So, a second definition of the term reliability is needed.

Let $Q_{\overline{T}} = (g_{n+1}, g_{n+2}, \ldots\ldots, g_n)$ be the set of modelled faults that are not detected by test sequence T. Let $NM = (g_{nm1}, g_{nm2}, \ldots\ldots, g_{nml})$ be the set of unmodelled faults. Certain physical defects like shorts can be modelled as logical faults while some physical defects like changes in voltage and loading can not be modelled as logical faults. Defects that can not be modelled may effect the performance of the digital system.

Definition 2: The reliability of a circuit associated with a test sequence T is given by

$$R_T = \left[\prod_{i=1}^{n} (1-ki \int_{to}^{t} fi \cdot dt)\right]\left[\prod_{j=n+1} (1 - k_j \int_{t_j}^{t} f_j dt)\right]\left[\prod_{\ell=nml}^{nm\ell} (1- \int_{t_o}^{t} f_l \cdot dt)\right] \quad (14)$$

where fi, fj, and fl are the probability density functions,

ki and kj are weighting constants as before and

t0 is the time the test T was applied last,

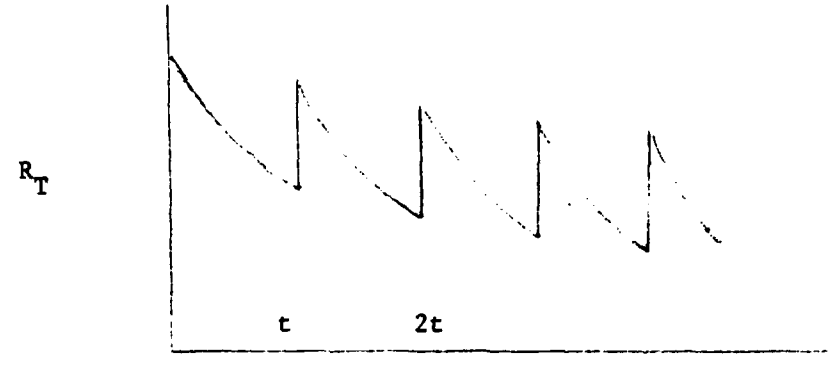
tj is the time when a test checked for fault gj and

$t_{-\alpha}$ is the time when the circuit was designated operative.

The first term in equation (14) is due to the faults (g1,g2,...,gn) that are detected by test T. The second term is due to the modelled faults that are not detected by test T and the third term is due to the unmodelled faults. The third term may be dropped if the unmodelled faults are not critical or if the failure rates of such faults are not readily available.

The general form of the reliability curve given by equation (14) is shown in Figure 7. When the reliability reaches a value which is not ac-

ceptable, the circuit must be replaced.



The first definition may be used when an user is interested only in the faults that are detected by his test set. But, if more accurate reliability measures are desired, the second definition should be used.

AN EXAMPLE

The reliability calculations are illustrated using the following circuit.

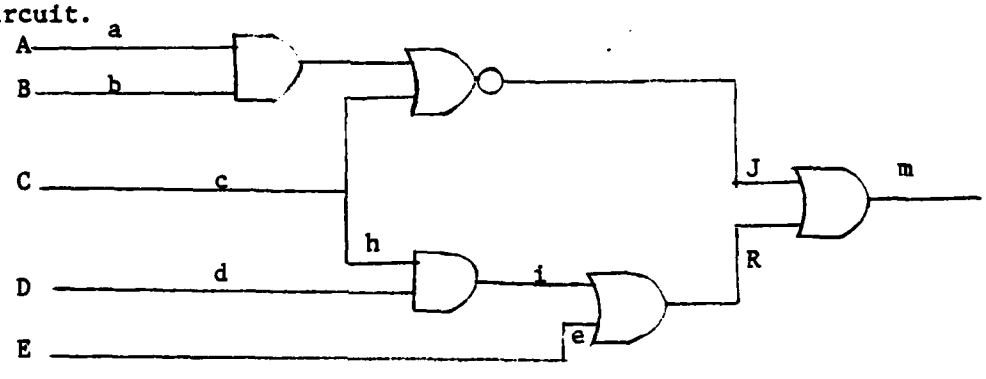


Figure 8. An Example

This circuit has 24 stuck faults under the single fault assumption. This number is reduced to 10 faults on input signals and 4 faults on branches at fanout points (g and h stuck faults) using fault collapsing. The table given below contains 6 input patterns that detect all faults in the circuit.

	Inputs	Faults Detected
	A B C D E	
1	1 1 0 1 0	$a_o, b_o, c_o, e_1, f_o, h_i, i_1, j_1, k_1, r_1, m_1$
2	0 1 0 1 0	$a_1, c_1, f_1, g_1, J_o, m_o$
3	1 0 0 1 0	$b_1, c_1, f_1, g_1, J_o, m_o$
4	0 x 1 1 0 x 0	$c_o, d_o, h_o, i_o, k_o, m_o$
5	x 0 1 0 0 0 x	$d_1, e1, g_o, i_1, J_1, k_1, m_1$
6	x x 1 0 1	e_o, k_o, m_o

Table - 1

Let $f_{a_o} = \lambda e^{-\lambda t}$ be the probability density function (pdf) of the fault a stuck at 0. For the sake of this example, we will assume that all faults have identical pdf's. We will assign a 1 to the weighting constants. Test set T = (1,3,5) detects 75% of the faults in the circuit.

Q_T = (a0,a1,b0,b1,c0,c1,d1,e1,f0,f1,g0,g1,h1,i1,j0,j1,k1,m0,m1)

Using the first definition,

$$R_T = \prod_{i=1}^{19} (1 - \int_o^t \lambda e^{-\lambda r} dr) = e^{-19\lambda t}$$

If the mean time between failures ($1/\lambda$) has a value of 5×10^4 hours, then

$$R_T = e^{-19x2x10^{-5}t}$$

For a minimum reliability of 0.9, the time between testings can be calculated:

$$t = \ln(R_T)/(-38 \times 10^{-5}) = 277 \text{ hours}$$

Thus, the test T should be applied every 277 hours in order to maintain a reliability of 0.9.

While using the second definition, we will ignore faults that can not be modelled. The set of faults that are not detected by test set T is

$$\bar{Q}_T = (d0, e0, h0, i0, k0).$$

Assuming that all faults were tested at time t=0,

$$R_T = [\prod_{i=1}^{19} (1 - \int_0^t \lambda e^{-\lambda r} dr)][\prod_{j=20}^{24} (1 - \lambda e^{-\lambda r} dr)]$$

When $\lambda = 2 \times 10^{-5}$, for a reliability of 0.9, the time between tests is calculated to be t = 219.5 hours.

Table (2) lists few other values for time between test for various reliabilities.

TABLE 2

R_A	MTBF	λ	t_1 with 1st def.	t_1 with 2nd def.
0.75	50.000 hrs.	2×10^{-5} hr	757 hrs	559 hrs
0.8	50,000	2×10^{-5}	587	464
0.9	50,000	2×10^{-5}	277	219.5
0.95	50,000	2×10^{-5}	134	107
0.75	90,000	1.14×10^{-5}	1328	1051
0.8	90,000	1.14×10^{-5}	1030	815
0.9	90,000	1.14×10^{-5}	486	385

Chapter III

ASSIGNMENT OF WEIGHTS TO FAULTS

In a microprocessor system or in any digital system, certain faults are more crucial to the operation of the system than others. For example, in most IC's, the enable and power lines are the most critical faults on these are fatal to the system. It may be useful to assign weights to the signals according to their importance in a circuit so that better reliability measures can be obtained. It is very difficult to come up with a procedure which assigns absolute values to weights since the importance of a signal not only depends on its function in a circuit, but the relative importance of signals may be biased by user views. Here we attempt to outline a heuristic procedure that enables an user to assign weights to various faults.

Faults in a digital system may be classified into four groups.

1. Very Important Faults: These faults are very critical to the operation of the system. It may be useful to include faults whose relative importance is not entirely clear. This leads to a conservative estimate of the circuit reliability, but the weights may be changed when the significance of faults becomes discernible. We can assign a weight of 1 to faults in this class.
2. Important Faults: There may be some faults whose presence impairs the operation of the circuit but does not create a hazardous outcome. These faults may become critical to the circuit operation or prolonged existence. For example, a stuck input in a shift register does not effect the output immediately, but if the fault remains it may lead to an erroneous results. Such faults are assigned a weight of 0.75.
3. Unimportant Faults: Faults in this class may not be critical to the operation of the circuit at all, even in continued existence. For example, in most computers, a fault in a clock circuit which changes the cycle time slightly

is not critical to the operation of the system. A weight of 0.5 is assigned to faults in this category.

4. Don't Care Fault: Some faults may not effect the operation of the circuit, especially if redundant logic is used. These faults can be ignored in some circumstances, and so a weight of 0.0 is assigned to such faults.

At this point we would like to draw an analogy from an automobile. Examples of very important faults include faults in the circuit that controls the brakes, faults in the ignition circuit. A fault in spedometer or a gasoline indicator can be critical in prolonged existence. A faulty spare tire can also be classified as an important fault. A failure in the heater or air conditioner operation can be classified as unimportant faults. A defect in accessories, such as the radio or clock, can be considered as a don't care faults.

BOUNDS ON TIME BETWEEN TESTS

In this section the upper and lower bounds on the time between tests (TBT) are derived. The reliability of a circuit associated with a test set S drops below the permissible minimum if the circuit is not exercised against the test set S at least once every TBT_{max} for that test. Application of a test more frequently than that indicated by TBT_{min} may not be necessary to maintain the desired reliability.

Let S1,S2,...,Sn be a number of test sequences such that

$$S_1 U S_2 U S_3 U \ldots U S_n = G$$
$$S_L \neq S_j \quad \text{if} \quad i \neq j \qquad (15)$$

Let us consider a test sequence S. Let g be the most important fault among the faults detected by S. Let t_s denote the time between test for test set S.

Upper Bound on t_s: The maximum value of t_s satisfies the following equation

$$1 - k_g \int_o^{t_s} f_g \cdot dt = R_{min} \tag{16}$$

where k_g is the weight assigned to fault g, fg is the pdf associated with g and R_{min} is the minimum acceptable reliability.

Lower Bound on t_s: Let $Q_s = (g1, g2, \ldots gm)$ be the set of faults detected by the test set S. Assuming that all the faults are equally important in the circuit, we can assign a weight of 1 to all faults in Q_s. Then, the minimum value of t_s satisfies

$$\prod_{i=1}^{m} (1 - \int_o^{t_s} f_{gi} \cdot dt) = R_{min} \tag{17}$$

The actual value of the time between tests for the test set S, t_s, depends on the real values of the weights k_i's. If reasonable weights can not be assigned to faults in a digital system, TBT_{min} should be used for scheduling the check-outs of the circuits.

FAILURE RATES OF VARIOUS FAULTS

Sometimes it may not be practical to calculate the failure rates of each and every fault in a digital system. In such cases, it is desirable to classify the faults into various groups depending on either the proximity of the nature of the failure mechanisms or on the closeness of the failure rates. Once similar faults are grouped together, identical failure rates for faults in a group may be assumed. Since the grouping of faults into various categories impacts the accuracy of the reliability measures, the classification should be conducted very carefully.

Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the failure rates respectively of the faults g1, g2,...,gn, the modelled faults that are detected by the test T. Let

$\lambda_{n+1}, \lambda_{n+2}, \ldots \lambda_r$ be the respective failure rates of the modelled faults $g_{n+1}, g_{n+2}, \ldots, g_n$ that are not detected by the test T while $\lambda_{nm1}, \lambda_{nm2}, \ldots$ $\lambda_{nm\ell}$ are the failure rates of unmodelled faults.

Assuming exponential failure rates, the sum of these failure rates gives the failure rate of the digital system under consideration.

$$\lambda_{ckt} = \sum_{i=1}^{n} \lambda i + \sum_{j=n+1}^{r} \lambda j + \sum_{\ell=nm1}^{nm\ell} \lambda \ell \qquad (18)$$

Since it is not possible to find the individual failure rates of unmodelled faults (or their nature), it is reasonable to conglomerate the failure rates of unmodelled faults into a single term λnm.

$$\lambda_{ckt} = \sum_{i=1}^{n} \lambda i + \sum_{j=n+1}^{i} \lambda j + \lambda nm \qquad (19)$$

The fault rates of modelled faults may similary be bundled into one or more groups depending on the similarity of the faults and the required accuracy of the model. We will illustrate bundling of faults for LSI circuits.

For LSI circuits it is observed (Tees [71], Kasouf [78]) that the failure rates of the pins (primary inputs and outputs) are higher than the failure rates of the gates. This leads to a classification of faults into 3 groups - faults on pins, faults on gates and unmodelled faults. The failure rate can be written as

$$\lambda_{LSI} = \lambda_{pins} + \lambda_{gates} + \lambda_{nm} \qquad (20)$$

where λ_{pins} is the total failure rate of faults on pins,

λ_{gates} is the failure rate due to faults on gates.

Equation (20) is similar to the equations in Tees [71] and Kasouf [78] which are of the general form

$$\lambda_{LSI} = C_1 \cdot P + C_2 \cdot G + C_3 \qquad (21)$$

where P is the number of pins and G is the number of gates in a LSI circuit. Constant C1, C2, and C3 are evaluated from the data supplied by the vendor from his own failure analysis and user feedback. In our case, the model is similar, but the constants reflect the failure rates of modelled logic faults on pins and gates. If the values of λ_{pins} and λ_{gates} in equation (20) can be obtained, failure rates of individual faults can be calculated assuming identical fault rates for faults in a group. These failure rates can then be used to calculate reliability of the circuit and the frequency of testing to achieve a desired reliability.

Available data about failures in semiconductor devices is typically of the form, n failures in m hours of operation. This type of data enables us to estimate the failure rate λ_{ckt} of the circuit as a whole. However, if, in addition to observing that a failure has occurred, the nature of the fault - a fault on a pin, a fault on a gate or else - is discovered, we would have data to estimate λ_{pin} and λ_{gates} in equation (20). If the bundling results in a different model, the failure rates due to faults in each group can similary be estimated.

Chapter IV

EXTENSIONS AND SOME OTHER APPLICATIONS

Although equation (20) is derived using the LSI failure mechanism, this model (equation (19)) can be used for any circuit. For a complex circuit using a large number of circuits, two extensions of the model are possible.

1.

$$\lambda_{Ckt} = \sum_i \lambda_{ICi} + \sum_j \lambda_{wire\ j} + \lambda_{nm} \tag{22}$$

where

$\lambda_{IC\ i}$ is the failure rate of IC i considered as one piece.

$\lambda_{wire\ j}$ is the failure rate of faults on the interconnecting wire j which depends on the wire length etc.,

λ_{nm} is the failure rate of all the other faults.

2. $$\lambda_{Ckt} = \lambda_{pins} + \lambda_{wires} + \lambda_{gates} + \lambda_{nm} \tag{23}$$

where

λpins is the combined failure rate of faults on the pins of all the IC's.

λwires is the combined failure rate due to the faults on interconnecting wires.

λgates is the combined failure rate of the faults on all gates.

λnm is the failure rate due to all the other faults.

As noted in the introduction, the model can be used to evaluate test sequences. Recently, a statistical method for test sequence evaluation is developed (case [16]). Generation and evaluation of test sequences have been the subjects of much investigation in the past. But here, test sequence

evaluation is a by-product.

In the previous sections, we have shown how to calculate the reliability associated with a test sequence. This reliability can be used as a measure to rank test sequences: a better test will have higher reliability and/or lower frequency of testing for a given confidence level.

One other possible application is to use our model in deriving reliability equations for redundant fault-tolerant systems. There have been many mathematical models developed for redundant systems like triple modular redundant (TMR), hybrid redundant, and stand-by spare. (See for example Bouricius [71], Mathur [71]). In almost all these studies, a simple exponential function is used to represent reliability of constituent systems. If the model presented in this report is used instead of simple exponential distribution, more accurate reliability predictions can be derived.

CONFIDENCE INTERVAL FOR THE RELIABILITY R_A

As the reliability measure R_A defined earlier is decreasing as time increases, we would like to maintain the reliability at some level by applying the test sequence A periodically. In order to find the associated confidence limit to maintain R_A at some given level and finally to present the time interval which is required to check the given circuit periodically, let's define some terminologies first.

Definition

Confidence level 100 β% ($0 \leq \beta \leq 1$) of the reliability is the percentage value of the probability such that we are at least 100 β% sure that the reliability is contained in some interval (R_L, R_u) which is called the 100 β% confidence interval of R_A while R_L and R_u are called lower and upper confidence limit of R_A respectively. i.e. Prob. $[R_L < R < R_u] = \beta$.

Remark: Frequently one is interested in one-sided confidence interval and the definition given above can be modified accordingly i.e. if one is interested in lower confidence limit only (as is the case with us) then Prob. $[R_L < R] = \beta$ is the associated probability statement and 100 β% is called the confidence level of R_A and R_L is called the (lower) confidence limit of R_A etc.

If identical failure rates are assumed among each of Q_A, $Q_{\bar{A}}$ and NM, the reliability R_A will eventually be a function of time t with parameter $\lambda = \lambda_1 + \lambda_2 + \lambda_3$ where λ_1: failure rate for modeled, detected by test seq. A.

λ_2: failure rate for modeled, undetected by test seq. A.

λ_3: failure rate for modeled, unmodeled faults.

i.e. $R_A = R_1(t; \lambda_1, \lambda_2, \lambda_3) = R_2(t; \lambda)$ where $\lambda = \lambda_1 + \lambda_2 + \lambda_3$. (23)

Further assume exponential failure model for each of Q_A, $Q_{\bar{A}}$ and NM, the para-

meters λ_1, λ_2 and λ_3 respectively and assume failures in each set $Q_A, Q_{\bar{A}}$ and NM occur independently of each other.

Here, one performs an experiment and observes a Poisson process in which r failures are observed in time $t_r = n\, t_o$, where t_o is the time of termination of the life test and n is the number of items in the sample subjected to life test (Type I Censoring See Mann [75]) and $r = r_1+r_2+r_3$, where r_1, r_2 and r_3 are the number of faults in $Q_A, Q_{\bar{A}}$ and NM respectively observed in time interval t_o.

Let K_i be the random variable of which r_i is the realization, i = 1,2,3 and let $K = K_1 + K_2 + K_3$. Since K_i's follow Poisson distribution with parameter $\lambda_i\, t_r$ (see Mann [75]), K also follows Poisson with parameter λt_r by the independency assumption of K_i's for i = 1, 2 and 3.

Let $\hat{\theta} = n\, t_o/r = t_r/r$ and using the identity Prob $(K \leq r) = \text{Prob}[\chi^2(2r+2) > 2\lambda\, t_r]$, where $\chi^2(\nu)$ is the random variable representing chi-square distribution with ν degrees of freedom, we can show 100 $(1-\alpha)$% upper confidence limit λ_U for λ is $\dfrac{\chi^2_{1-\alpha}\,(2r+2)}{2r\,\hat{\theta}}$, where $\chi^2_q\,(\nu)$ means q-th quantile of chisquare distribution with ν degrees of freedom. Here, one can obtain the lower confidence limit R_L to maintain R_A at confidence level $100(1-\alpha)$%, since R_A is a decreasing function of γ . i.e.

$$R_L = R_2(t;\lambda_U) = R_2\; t;\; \frac{\chi^2_{1-\alpha}(2r+2)}{2r\,\hat{\theta}}\,, \tag{24}$$

where R_2 is given in (23)

Further, to find the time period t_o^* required to periodic checking procedure to maintain the reliability at some given level R_{min}, we can simply use (24) i.e., t_o^* is the value of t which satisfies the equation

$$R_{min} = R_2\;\; t;\; \frac{\chi^2_{1-\alpha}(2r+2)}{2r\,\hat{\theta}} \tag{25}$$

i.e. if we check the given circuit at t_o^* time period, we can at least be 100 $(1-\alpha)$% sure that the reliability R_A is at least R_{min}.

Note: Since we do not know the true value of R_T, which is a function of unknown λ_i's, it is more realistic to estimate the lower bound of R_T instead of "mean" value of R_T and use the time interval to maintain the lower bound of R_T. In other words, results in this section is more conservative (more frequent checking) compared to the case in the examples in the appendix at the end of Chapter II.

Chapter V

Point and Interval Estimation for Life Testing Procedure of LSI/MSI Reliability Models - System Point of View

In theoretical studies of equipment reliability, one is often concerned with systems consisting of many components, each subject to an individual pattern of malfunction and replacement, and all parts together making up the failure pattern of the equipment as a whole.

Drenick [1960] showed that a complex piece of equipment, after an extended period of operation, will tend to exhibit a failure pattern with an expoential distribution for the time between failure (inter-arrival times) and that the time up to the first failure is also nearly exponentially distributed. Hence LSI/MSI reliability model can be described by expoential density function

$$f(t) = (1/\theta)\exp(-t/\theta) \qquad (1)$$

where θ is the average life time and t is the time to first failure of LSI/MSI.

One is often interested in estimating θ or, if mission time t_m is specified, in estimating reliability of an LSI/MSI component $R(t_m) = 1 - F(t_m)$ where $F(t)$ is the cumulative distribution function for $f(t)$. i.e. $R(t_m)$ is the probability of no fault occurs in the given LSI/MSI until time t_m.

In the following notes, the estimation procedure of θ and $R(\cdot)$ will be explained and demonstrated through examples for the various sampling situations. Moreover the procedure of computing the time period to maintain the given LSI/MSI component at the specified reliability level (lower confidence limit for $R(\cdot)$) is illustrated.

1. Testing without replacement

In the life testing procedure of LSI/MSI, a faulty LSI/MSI is not replaced until the testing is terminated.

(a) Estimation for Type II censoring

Suppose that a sample of size n LSI(MSI) with failure time distribution given by (1) has been subjected to life testing and that the test is terminated at the time of the r^{th} failure, $r \leq n$. More generally, one might assume that a sample of size n has been randomly selected from a one-parameter exponential population, and that experimentation is terminated at the time that r^{th} observation becomes available.

The joint density function of the ordered observations $X_{(1)},\ldots X_{(r)}$, $X_{(i)} \leq X_{(i+1)}$, i=1,...,r-1 and r=1,2,...,n is given by

$$f_{X_{(1)},\ldots,X_{(r)}}(x_1,\ldots,x_r) = \frac{n!}{(n-r)!\theta^r} \exp\left\{ -\frac{\sum_{i=1}^{r} x_i + (n-r)x_r}{\theta} \right\}$$

$$0 \leq x_i \ldots \leq x_r$$

Clearly, if the right-hand side of this is maximized with respect to θ, one obtains, as the maximum - likelihood estimator (MLE) of θ,

$$\hat{\theta} = \frac{\sum_{i=1}^{r} X_{(i)} + (n-r)X(r)}{r}$$

Epstein (1953-1954) showed that $\hat{\theta}$ is unbiased, sufficient and complete for θ and because $\hat{\theta}$ is based on a complete sufficient statistic, it is, by the Lehmann-Scheffe'-Blackwell theorem, unique minimum-variance estimator (UMVUE) among unbiased estimator of θ.

Also $\mathrm{Var}(\hat{\theta}) = \theta^2/r$.

If we let $S_i = (n-i+1)(X_{(i)} - X_{(i-1)})/\theta$, i=1,...r, S_i has an independent expoential distribution with scale parameter equal to 1 and it is well known that $2S_i$, i=1,...r, has an independent chi-square distribution with 2 degrees of freedom.

Also one can easily infer that $\sum_{i=1}^{r} 2S_i = 2r\hat{\theta}/\theta$ (being the sum of r independent chi-squares) has a chi-square distribution with 2r degrees of freedom. A lower confidence bound on θ at confidence level $1-\alpha$ is therefore given by $2r\hat{\theta}/x^2_{1-\alpha}(2r)$, where $x^2_\alpha(k)$ is the 100 th percentile of a chi-square distribution with k degrees of freedom. The interval $[2r\hat{\theta}/x^2_{1-\alpha/2}(2r),\ 2r\hat{\theta}/x^2_{\alpha/2}(2r)]$ is a two-sided confidence interval for θ at confidence level $1-\alpha$. To obtain such intervals one simply substitutes calculated values for $\hat{\theta}$ and tabulated values of chi-square percentile (see example below). The minimum variance unbiased point estimator of $R(t_m)$ for r=n is [See Pugh (1963) and Basu (1964)]

$$R^*(t_m) = \begin{cases} (1 - \frac{t_m}{n\hat{\theta}})^{n-1}, & n\hat{\theta} > t_m, \\ 0, & n\hat{\theta} \leq t_m. \end{cases}$$

A lower confidence bound for $R(t_m)$ at level $1-\alpha$ can be obtained by substituting $2r\hat{\theta}/x^2_{1-\alpha}(2r)$ for θ in the expression $\exp(-t_{m/\theta})$. Two-sided confidence intervals are obtained similarly.

These results are summarized in Table I below.

The time period t_m^* to maintain the reliability at a given level, say β, can be obtained by equating β with the lower confidence limit for $R(t_m)$ and solving for t_m.

Table I

parameter	point estimate	interval estimate
θ	$\frac{\sum_{i=1}^{r} x_{(i)} + (n-r)\, X_{(r)}}{r}$	$\left[\frac{2r\hat{\theta}}{x^2_{1-\alpha/2}(2r)}, \frac{2r\hat{\theta}}{x^2_{\alpha/2}(2r)}\right]$
$R(t_m)$	$(1 - \frac{t_m}{n\hat{\theta}})^{n-1}$, $n\hat{\theta} > t_m$; 0, $n\hat{\theta} \leq t_m$	$\exp(-t_m/[\frac{2r\hat{\theta}}{x^2_{1-\alpha}(2r)}])$ (lower confidence limit)

Example 1.

Let n=24 LSI(MSI)'s are put in life testing and the test was terminated at r=3rd failure.

If the observed exponential failure times in hours from the above censored samples of size 24 are 6200,9200 and 16900 then $\hat{\theta}=[6.2 + 9.2 + 22(16.9)]\cdot 10^3/3 = 129,100$ hours. 95% two-sided confidence interval ($\alpha=0.05$) is [53605.5, 626040.6]. If mission time t_m = 5,000 hours, lower confidence limit for $R(t_m)$ at level .95 ($\alpha=.05$) with t_m=5,000 hours is $\exp\left(-5000/\left[\frac{6x129100}{12.59}\right]\right)=.92$ i.e. we can be at least 95% sure that the reliability is at least .92. To obtain the time period to maintain the reliability at least $\beta=.95$ level, we solve

$$.95 = \exp\left[-t_m \Big/ \frac{6x129100}{12.59}\right] \text{ and get}$$

$t_m^* = 3155.8$ i.e. we might have to check the LSI after using 3155 hours periodically to be 95% sure that the reliability is at least .95.

(b) Estimation for Type I censoring

If a life test is terminated at a specified time t_o, the joint p.d.f. of $x_{(1)},\ldots,x_{(r)}$ is given by

$$f_{x_{(1)},\ldots,x_{(r)}}(x_1,\ldots,x_r) = \frac{n!}{(n-r)!\theta^r}\exp\left\{-\left[\sum_{i=1}^{r} x_i + (n-r)t_o\right]\right\},$$

$$o\le x_1\le\ldots\le x_r\le t_o.$$

The maximum likelihood estimator of θ is easily seen to be

$$\hat{\theta} = \frac{\sum_{i=1}^{r} x(i) + (n-r)t_o}{r}, \qquad r \neq 0.$$

This estimator is biased for small samples but has all the desirable asymptotic properties associated with the MLE of θ under Type II censoring.

To obtain a confidence interval for θ, one may consider the probability that the number of failures, K, in a sample of size n is equal to r at time t_o:

$$P(K=r) = \frac{n!}{(n-r)!r!} [R(t_o)]^{n-r}[1-R(t_o)]^r.$$

Then a conservative (because K is discrete and the failure times are not used) lower confidence bound δ at level $1-\alpha$ for $R(t_o)$ is given by the solution of

$$\sum_{i=n-r}^{n} \binom{n}{i} \delta^i (1-\delta)^{n-1} = \alpha = \int_o^\delta \frac{\Gamma(n+1)}{\Gamma(n-r)\Gamma(r+1)} x^{n-r+1}(1-x)^r dx$$

From the relationship between the binomial and the beta distributions, it can be seen that δ is equal to

$1 - V_{1-\alpha}(r+1, n-r) = V_\alpha(n-r, r+1)$, the 100$\alpha$th percentile of the beta distribution with parameters n-r and r+1. From this, one can determine, from $R(t_o)=\exp(-t_o/\theta)$, a lower confidence bound for θ at level $1-\alpha$ as $t_o/\ln(1/\delta)$.

Table II

parameter	point estimate	interval estimate
θ	$\hat{\theta} = \frac{\sum_{i=1}^{r} X_{(i)} + (n-r)t_o}{r}$, $r \neq 0$	$t_o/\ln(1/\delta)$ (lower confidence bound)
$R(t_o)$	$R(t;\hat{\theta})$	$\delta = V_\alpha(n-r, r+1)$ (lower)

Example 2.

Sample size n=15 LSI's were put in life testing and the test was terminated at t_o=240 days. The number of failures up to time t_o were r=2; then a 90 % lower confidence bound for $R(t_o)$ is .6827 and a 90 to lower confidence bound for θ is $240/\ln(1/.6827) \doteq 630$ days.

2. Testing with replacement.

(a) Estimation for Type II censoring

If a life test is performed with replacement and testing is terminated at the time of the rth observed failure, then the joint p.d.f. of the ordered failure times $X_{(1)},\ldots,X_{(r)}$ is given by

$$f_{x_{(1)},\ldots,x_{(r)}}(x_1,\ldots x_r)=(n\lambda)^r\exp(-n\lambda x_r),0<x_1\leq\ldots\leq x_r,$$

where $\lambda = 1/\theta$.

From this expression one can obtain the m.L.E. of λ as $(r/X_{(r)})/n$ and of θ as $(nX_{(r)})/r$. For this model, one can also write the joint p.d.f. of $x_{(1)},\ldots,x_{(r)}$ as

$$(n\lambda)^r\exp[-n\lambda\sum_{i=1}^{r}(x_i-x_{i-1})],\ x_o\equiv 0<x_1\leq\ldots\leq x_r$$

In other words, each $Si = X_i-X_{i-1}$, $i=1,\ldots,r$, with $X\equiv 0$, has an independent exponential distribution with scale parameter $(\lambda n)^{-1}$. Since $X_{(r)} = \sum_{i=1}^{r} Si$, the density of $X_{(r)}$ for testing with replacement is given by

$$f_{x_{(r)}}(x_r) = \frac{(n\lambda)^r x_r^{n-1}}{(r-1)!}\exp(-n\lambda x_r).$$

Consequently, $2n\lambda X_{(r)}$ has a chi-square distribution with 2r degrees of freedom, and confidence bounds for λ, θ or reliability can be obtained as functions of observed values $X_{(r)}$ and values to be obtained from tables of the chi-square distribution.

Table III

parameter	point	interval
θ	$\hat{\theta} = nX_{(r)}/r$	$\left[\frac{2nX(r)}{X^2_{1-\frac{\alpha}{2}}(2r)} \quad \frac{2nX(r)}{X^2_{\frac{\alpha}{2}}(2r)}\right]$
R(t)	$R(t;\hat{\theta})$	$R(t; \frac{2nX(r)}{X^2_{1-\alpha}(2r)})$ (lower limit)

Example 3

For a sample (with replacement) of size 5 LSI's from an exponential population for which the first observed failure time is 5600 hours and the fourth and last observed time, occuring at the termination of the life test, is 20,000 hours, the MLE for θ is $5(20,000)/4=25,000$ hours and an 80 % lower confidence bound for θ is $10(20,000)/X^2_{80}(8)\cong 18,100$ hours.

(b) Estimation for Type I censoring.

In this case, one observe a Poisson process in which r failures are observed in time $t_r=nt_o$, where t_o is the time of termination of the life test and n is the number of items in the sample subjected to life test. For this model, the number of observed failures, K, is a random variate with

$$P(K=r)=(\frac{1}{r!})(\lambda t_r)^r \exp(-\lambda t_r) \text{ with } \lambda=1/\theta.$$

The M.L.E $\hat{\lambda}$ of λ is $r/t_r = r/nt_o$ and the MLE $\hat{\theta}$ of θ is $\hat{\lambda}^{-1}=nt_o/r$.

then

$$P(K\leq r)= \sum_{k=0}^{r} \frac{1}{k!} (\lambda t_r)^k \exp(-\lambda t_r)$$

$$= \int_{\lambda t_r}^{\infty} \frac{z^r}{r!} \exp(-z)dz,$$

An incomplete gamma function with shape parameter r + 1. Therefore Prob($K \leq r$) = Prob $[x^2(2r+2) > 2\lambda t_r]$. Thus if one observes r failures, then, with probability at least 1-α, $2\lambda r\hat{\theta} < \chi^2_{1-\alpha}(2r+2)$, or $2nt_o/x^2_{1-\alpha}(2r+2)$ is a conservative lower confidence bound for θ at confidence level 1-α. Similarly, it can be shown that

Prob $(K>r)$ = Prob$[\chi^2(2r)\ 2\lambda t_r]$.

Hence a conservative upper confidence bound for θ at level 1-α is given by $2r\hat{\theta}/x^2_{\alpha}(2r)$, and a conservative two-sided confidence interval at level 1-α is

$$\left[\frac{2r\hat{\theta}}{\chi^2_{1-\alpha/2}(2r+2)}\ ,\ \frac{2r\hat{\theta}}{\chi^2_{\alpha/2}(2r)}\right]$$

Cox (1953) discusses the interval

$$\left[\frac{2r\hat{\theta}}{\chi^2_{1-\alpha/2}(2r+1)}\ ,\ \frac{2r\hat{\theta}}{\chi^2_{\alpha/2}(2r+1)}\right]$$

which he states is slightly narrower than that given above but sometimes has a true confidence coefficient less than 1-α.

Table IV

parameter	point	interval
θ	$\hat{\theta}=nt_o/r$	$\left[\frac{2r\hat{\theta}}{\chi^2_{1-\alpha/2}(2r+2)}\ ,\ \frac{2r\hat{\theta}}{\chi^2_{\alpha/2}(2r)}\right]$
R(t)	$\hat{R}(t) = R(t;\hat{\theta})$	$R\left(t;\ \frac{2r\hat{\theta}}{\chi^2_{1-\alpha}(2r+2)}\right)$ (lower limit)

Note that $R(t_i\theta)$ is an increasing function of θ.

Example 4

If 200 LSI's are life tested with replacement for 1,000 days and failures are observed, a 90% lower confidence bound for the mean time to failure is 400 (1,000)/$\chi^2_{.90}(8)$=400,000/13.36 = 29,900 days. As an example in real life, we illustrate the estimation procedure using the data from G. Kasouf and S. Mercurio (1978).

Certified data from an airborne radar processing systems (RPS) operating in a simulated airborne inhabited environment is used to calculate the observed LSI/MSI failure rate. The certified data was accumulated during the RPS 1600-hour test-analyze-and-fix program followed by a 351-hour fixed length reliability demonstration test. Analysis of data shows an accumulation of 3.9 and 4.6 million operating hours for LSI and MSI circuits, respectively. One LSI and no MSI failures were experienced. In this case t_r= (3.9 + 4.6) x 10^6 = 8.5 x 10^6 and $\hat{\lambda}$=1/(8.5x10^6)=0.118/10^6 hrs and r=1. 60% (α=.4) confidence interval for λ(failure rate) is

$$\left[\frac{\chi^2_{.2}(2)\hat{\lambda}}{2}\ ,\ \frac{\chi^2_{.8}(4)}{2}\hat{\lambda}\right] = [.026314\ ,\ .35341]$$

95% lower confidence bound for θ is

$$\frac{2r\hat{\theta}}{\chi^2_{.95}(4)} = \frac{2 \times \frac{10^6}{0.118}}{9.488} = 1.78637 \times 10^6 \text{ hours}$$

To obtain the time period to maintain the LSI at B=.98 level, we solve

$$.98 = \exp[-t/(1.78637 \times 10^6)]$$

and get

$$t^*_m = 3.609 \times 10^4 \text{ hours}$$

i.e. we may have to check the LSI's in 3.6 x 10^4 hours period to be 95% sure that the reliability is at least .98.

Appendix A

STATISTICAL PROPERTIES OF QUASI-RANGE IN SMALL SAMPLES FROM A GAMMA DENSITY

Introduction

Karl Pearson [1920], studied combinations of order statistics which were later named quasi-ranges by Mosteller [1946]. Quasi-ranges are simple to obtain and, for moderate sample sizes, yield more efficient estimators of standard deviation than do sample ranges. In cases where abundant data are available and where the cost of complicated data reduction far outweighs the cost of sampling, the use of quasi-range in estimation problems is satisfactory.

Quasi-range may be defined as follows. Consider the ordered sample values $Y_1, Y_2, \ldots, Y_n$ where $a < Y_1 < Y_2 < \ldots < Y_n < b$. If the r smallest and r largest of these values are deleted, the range of the remaining (n-2r) values is defined to be the rth quasi-range, W_r. Symbolically,

$$W_r = Y_{n-r} - Y_{r+1},\ n > 2r + 1,\ b - a > W_r > 0. \qquad (1.1)$$

Note that sample range is simply the quasi-range, W_o.

Many authors have found quasi-ranges to be of particular interest and have developed numerous applications for them. Among others, Harter [1959] , Cadwell [1953], Chu [1957] , Benson [1949] and Ghosal [1957] have derived estimators based on quasi-range, and Rider [1959] has studied various quasi-range distribution.

As a natural extension of research first published by Gupta [1960] regarding the distribution of order statistics from a gamma density, this study of quasi-range is intended to provide further information needed by the analyst in applying the methods of simple estimation and inference that have been developed. It should also be noted that Prescott [1974] has given variances and covariances of order statistics from the gamma distribution. Though Tables I, II and III show the moments and quantiles of quasi-range for limited sample sizes and parameter values (n=1(1) 10 and α = 0(1) 2) only, methods described here can be easily extended.

Distribution of Quasi-Range in Samples From A Gamma Density

The density function of the standard gamma distribution with parameter α is

$$f(y) = (\alpha!)^{-1} e^{-y} y^{\alpha} \quad (y \geq 0), \tag{2.1}$$

where α is a non-negative integer. The corresponding cumulative distribution function may be expressed as a partial sum of Poisson probabilities:

$$F(y) = 1 - \sum_{i=0}^{\alpha} \frac{e^{-y} y^{i}}{i!} \quad (y \geq 0). \tag{2.2}$$

Let W_r be the rth quasi-range from a sample of size n with distribution F(y). It is well known (Harter [1959]) that the distribution of W_r can be expressed in integral form as

$$\phi(w_r) = \frac{n!}{r!(n-2r-2)!r!} \int_a^b [F(y)]^r [F(y+w_r) - F(y)]^{n-2r-2} \cdot [1-F(y+w_r)]^r f(y) f(y+w_r)\, dy \quad \text{for } 0 < w_r < b - a. \tag{2.3}$$

Substitute (2.2) into (2.3) and let $a_m(\alpha,r)$ be the coefficient of t^m in the expansion of $\left(\sum_{j=0}^{\alpha} t^j/j!\right)^r$ (See Prescott [1974]), then (2.3) can be written as

$$\begin{aligned}\phi(w_r) = {} & \frac{n!}{(r!)^2(n-2r-2)!(\alpha!)^2} \sum_{i_1=0}^{r} \binom{r}{i_1} (-1)^{i_1} \sum_{i_1=0}^{\alpha \cdot i_1} a_{i_2}(\alpha, i_1) \\ & \cdot \sum_{i_3=0}^{n-2r-2} \binom{n-2r-2}{i_3} (-1)^{i_3} \sum_{i_4=0}^{\alpha \cdot i_3} a_{i_4}(\alpha, i_3) \sum_{i_5=0}^{i_4} \binom{i_4}{i_5} \\ & \cdot \sum_{i_6=0}^{\alpha(n-2r-2-i_3)} a_{i_6}(\alpha, n-2r-2-i_3) \sum_{i_7=0}^{\alpha \cdot r} a_{i_7}(\alpha, r) \sum_{i_8=0}^{i_7} \binom{i_7}{i_8} \\ & \cdot \sum_{i_9=0}^{\alpha} \binom{\alpha}{i_9} \frac{w_r^{(i_4+i_7+\alpha-i_5-i_8-i_9)} \cdot e^{-w_r(i_3+r+1)}}{(i_1+n-r)^{(i_2+i_5+i_6+i_8+i_9+\alpha+1)}} \\ & \cdot (i_2+i_5+i_6+i_8+i_9+\alpha)!, w_r > 0. \end{aligned} \tag{2.4}$$

In order that this expression for the density of W_r be of use, it can be reduced to the following form,

$$\phi(w_r) = \sum_{j=0}^{n-2r-2} e^{-w_r(r+j+1)} \sum_{k=0}^{\alpha(j+r+1)} A(j,k) w_r^k, \; w_r > 0, \qquad (2.5)$$

where $A(j,k)$ is the accumulation of all coefficients in (2.4) involving the like powers of w_r and e. This function, which we call the *coefficient matrix* will be used as a notational convenience throughout this paper. Although it is impossible to express $A(j,k)$ in closed form, we can compute its numerical value in the summation process of (2.4) by evaluating the coefficient of $[\exp\{-w_r(r+j+1)\}]\, w_r^k$ where for specified values of n, r and α

$k = i_4 + i_7 + \alpha - i_5 - i_8 - i_9$.

Using $A(j,k)$, the moment generating function M(t) associated with quasi-range can be written as

$$M(t) = \sum_{j=0}^{n-2r-2} \sum_{k=0}^{\alpha(j+r+1)} \frac{A(j,k)\cdot k!}{(r+j+1-t)^{k+1}}, \; t < r + \alpha + 1, \qquad (2.6)$$

from which it is easily seen that the ith moment about the origin of the rth quasi-range is

$$\mu_i'(n,r,\alpha) = \sum_{j=0}^{n-2r-2} \sum_{k=0}^{\alpha(j+r+1)} \frac{A(j,k)(k+i)!}{(r+j+1)^{k+1+i}}. \qquad (2.7)$$

These values along with the ith central moments are tabulated in Table I.

The cumulative distribution function for W_r can be obtained from (2.5) and it may be transformed to a partial sum of Poisson probabilities as

$$\phi(w_r) = \sum_{j=0}^{n-2r-2} \sum_{k=0}^{\alpha(r+j+1)} \frac{A(j,k)\,k!}{(r+j+1)^{k+1}} \left\{ 1 - \sum_{i=0}^{k} \frac{[\exp\{-w_r(r+j+1)\}]\{w_r(r+j+1)\}^i}{i!} \right\} \qquad (2.8)$$

from which the p-quantile ξ_p of the density (2.5) can be calculated iteratively. A suitable iterative procedure which is quite efficient in computing ξ_p is given by

$$\xi_i = \xi_{i-1} - (p_{i-1} - p)(\xi_{i-1} - \xi_{i-2})/(p_{i-1} - p_{i-2}), \qquad (2.9)$$

where $\phi(\xi_p) = p$ and $\phi(\xi_i) = p_i$ at the ith iteration. To initiate the process, fairly arbitrary values of ξ_1 and ξ_2 may be chosen and using (2.8) p_1 and p_2 can be obtained. These values are tabulated in Table III.

Analysis of Results

Ghosal [1957] has published results of an analysis of quasi-range distributions associated with samples from an exponential density. It is well known that the gamma distribution reduces to the exponential distribution when $\alpha = 0$: hence a limited comparison between Ghosal's data and that of this study is possible. Table IV presents this comparison. It is seen that $\mu_1'(n,r,\alpha)$ and $\mu_2(n,r,\alpha)$ from Table I agree exactly with Ghosal's values, but values from Table II show some disagreement. Small discrepancies are revealed in the values for γ_1, and still larger errors appear for γ_2. To find the cause of these differences, Ghosal's formula was reapplied with more accuracy in each calculation. Exact agreement between data sets was obtained by this procedure, leading to the conclusion that Table I and II are correct for $\alpha = 0$.

Gupta [1960] has published tables of central and non-central moments for the distribution of order statistics in samples from a gamma density. Thus, it is possible to form a direct comparison between his data and that of this study. First moments, calculated from Gupta's results, were subtracted from those of Table I with the remainders written as absolute deviations. Results of this comparison are presented in Table V. A maximum deviation between the two data sets exists when $n = 10$, $\alpha = 2$, and $r = 3$. This discrepancy, amounting to 0.00079, reflects a deviation

from Gupta's results of 0.06 percent and represents the largest percent discrepancy found. Unfortunately, no recommendation based on this analysis can be made regarding the relative merit of either data set. It can only be surmised that discrepancies may be due solely to machine error computation (See Prescott [1974]). Higher moments are not compared since Gupta's work extends only to the derivation of formulas for the covariance between Y_i and Y_j. Although these formulas may be used to calculate quasi-range variance, the amount of work required to obtain a comparison is considered prohibitive. To obtain still higher moments, original deviations for covariance between powers of Y_{n-r} and Y_{r+1} are required, a task beyond the scope of this study. To test the behavior of error in moments of higher order, an alternative computer program was devised that differed from the original only in methods used to calculate coefficients A(j,k) in equation (2.5). Slightly different values for $\mu_i'(n,r,\alpha)$ were computed, but in general, the discrepancies tend to zero as the order of the moment increased. Two examples of this phenomenon appear in Table VI.

Another independent test was performed to determine the extent to which machine error accumulation contributes to overall error. The basic machine program used for all computation was revised so as to perform calculations in double precision arithmetic; only small differences less than 0.0001 were detected while computation time increased by at least a factor of two.

In addition to normal machine error, values in Table III contain a controlled error introduced by the iteration procedure tolerance limit. Thus a maximum error of two percent can exist in the lower tails of the distribution, since the quoted values of $\phi(w_r)$ may contain as much as 0.0001 error.

TABLE I

MOMENTS OF THE QUASI-RANGE DENSITY

$\alpha = 0$

n	r	μ'_1	μ'_2	μ'_3	μ'_4	μ_2	μ_3	μ_4
2	0	1.00000	2.00000	6.00000	24.0000	1.00000	2.00000	9.00000
3	0	1.50000	3.50000	11.2500	46.5000	1.25000	2.25000	11.0625
4	0	1.83333	4.72222	15.9722	67.7963	1.36111	2.32407	12.0069
	1	0.50000	0.50000	0.75000	1.50000	0.25000	0.25000	0.56250
5	0	2.08333	5.76389	20.2951	88.0914	1.42361	2.35532	12.5525
	1	0.83333	1.05556	1.80556	3.90741	0.36111	0.32407	0.84028
6	0	2.28333	6.67722	24.3015	107.533	1.46361	2.37132	12.9086
	1	1.08333	1.59722	3.00347	6.91088	0.42361	0.35532	1.01085
	2	0.33333	0.22222	0.22222	0.29630	0.11111	0.07407	0.11111
7	0	2.45000	7.49389	28.0484	126.232	1.49139	2.38058	13.1595
	1	1.28333	2.11056	4.26981	10.3267	0.46361	0.37132	1.12692
	2	0.58333	0.51389	0.60764	0.90394	0.17361	0.10532	0.18793
8	0	2.59286	8.23470	31.5776	144.276	1.51160	2.38641	13.3458
	1	1.45000	2.59389	5.56675	14.0379	0.49139	0.38058	1.2113
	2	0.78333	0.82722	1.10397	1.78711	0.21361	0.12132	0.24400
	3	0.25000	0.12500	0.09375	0.09375	0.06250	0.03125	0.03516
9	0	2.71786	8.91417	34.9204	161.736	1.52742	2.39032	13.4898
	1	1.59286	3.04699	6.87346	17.9656	0.51130	0.38642	1.27505
	2	0.95000	1.14389	1.67592	2.90439	0.24139	0.13058	0.28654
	3	0.45000	0.30500	0.27675	0.31515	0.10250	0.04725	0.06456
10	0	2.82897	9.54283	38.1013	178.670	1.53977	2.39306	13.6043
	1	1.71785	3.47845	8.17788	22.0545	0.52744	0.39030	1.32531
	2	1.09287	1.45614	2.29998	4.21866	0.26178	0.13643	0.31982
	3	0.61666	0.51055	0.53203	0.66983	0.13028	0.05651	0.08858
	4	0.20000	0.08000	0.04800	0.03840	0.04000	0.01600	0.01440

TABLE I

MOMENTS OF THE QUASI-RANGE DENSITY

$\alpha = 1$

n	r	μ'_1	μ'_2	μ'_3	μ'_4	μ_2	μ_3	μ_4
2	0	1.50000	4.00000	15.0000	72.0000	1.75000	3.75000	20.8125
3	0	2.25000	7.13889	29.2500	147.333	2.07639	3.84375	24.0404
4	0	2.74248	9.70216	42.3483	221.445	2.18098	3.77775	25.0139
	1	0.77257	1.09722	2.21094	5.70833	0.50036	0.59013	1.73654
5	0	3.10619	11.8704	54.3890	293.197	2.22195	3.71378	25.3323
	1	1.28762	2.34649	5.48081	15.5615	0.68853	0.68630	2.42653
6	0	3.39320	13.7526	65.5211	362.346	2.23880	3.66255	25.4053
	1	1.67117	3.57071	9.26752	28.3012	0.77789	0.70027	2.75526
	2	0.52051	0.50762	0.70830	1.27101	0.23669	0.19768	0.40128
7	0	3.62951	15.4183	75.8791	428.936	2.24497	3.62173	25.3746
	1	1.97529	4.72806	13.2999	43.0566	0.82631	0.69622	2.98760
	2	0.91088	1.18409	1.97755	4.01206	0.35438	0.25338	0.63624
8	0	3.82996	16.9145	85.5740	493.108	2.24592	3.58855	25.2978
	1	2.22641	5.81167	17.4336	59.2309	0.85479	0.68824	3.10855
	2	1.22193	1.91420	3.63859	8.11144	0.42110	0.27048	0.78760
	3	0.39248	0.29211	0.31307	0.43306	0.13807	0.09004	0.14036
9	0	4.00373	18.2741	94.6958	555.032	2.24420	3.56093	25.2013
	1	2.43976	6.82490	21.5883	76.4098	0.87246	0.67993	3.18317
	2	1.47967	2.65170	5.56665	13.3830	0.46229	0.27497	0.88927
	3	0.70647	0.71717	0.93851	1.49405	0.21806	0.12374	0.24225
10	0	4.15673	19.5212	103.318	614.875	2.24107	3.53748	25.0974
	1	2.62493	7.77397	25.7180	94.2998	0.88370	0.67258	3.22851
	2	1.69910	3.37648	7.67506	19.6406	0.48953	0.27455	0.96067
	3	0.96747	1.20447	1.82215	3.23287	0.26846	0.13740	0.31734
	4	0.31500	0.18973	0.16538	0.18663	0.09050	0.04860	0.06167

TABLE I

MOMENTS OF THE QUASI-RANGE DENSITY

$a = 2$

n	r	μ'_1	μ'_2	μ'_3	μ'_4	μ_2	μ_3	μ_4
2	0	1.87500	6.00000	26.2500	144.000	2.48437	5.68359	36.6086
3	0	2.81250	10.7878	52.0312	301.998	2.87765	5.50374	40.9339
4	0	3.42544	14.7036	75.9551	460.204	2.96999	5.24149	41.6138
	1	0.97367	1.69782	4.11572	12.5749	0.74978	1.00251	3.50675
5	0	3.87611	18.0098	98.0076	614.614	2.98563	5.05440	41.3822
	1	1.62279	3.64635	10.3094	34.8763	1.01290	1.10472	4.76581
6	0	4.23027	20.8718	118.398	763.998	2.97664	4.92076	40.8910
	1	2.10528	5.55944	17.5344	64.0997	1.12722	1.08395	5.35039
	2	0.65779	0.79518	1.35360	2.92363	0.36250	0.35364	0.86463
7	0	4.52082	23.3970	137.351	908.102	2.95919	4.82096	40.3414
	1	2.48696	7.36737	25.2509	98.1885	1.18240	1.04742	5.63643
	2	1.15112	1.85785	3.80536	9.33924	0.53477	0.43328	1.33669
8	0	4.76646	25.6583	155.061	1047.05	2.93914	4.74326	39.8006
	1	2.80134	9.05773	33.1679	135.698	1.21023	1.01360	5.77345
	2	1.54378	3.01072	7.03112	19.0292	0.62747	0.44589	1.62337
	3	0.49668	0.46066	0.60724	1.02141	0.21397	0.16589	0.31427
9	0	4.97880	27.7073	171.694	1181.09	2.91884	4.68059	39.2926
	1	3.06784	10.6355	41.1229	175.614	1.22387	0.98565	5.82849
	2	1.86855	4.17317	10.7850	31.5621	0.68170	0.43961	1.80521
	3	0.89411	1.13315	1.82935	3.55457	0.33372	0.21943	0.52999
10	0	5.16550	29.5816	187.384	1310.56	2.89921	4.62875	38.8224
	1	3.29847	12.1102	49.0226	217.213	1.23028	0.96132	5.84492
	2	2.14494	5.31528	14.8951	46.4916	0.71450	0.42902	1.92008
	3	1.22357	1.90430	3.56196	7.73555	0.40719	0.23548	0.68407
	4	0.39892	0.30044	0.32382	0.44742	0.14130	0.09123	0.14160

TABLE II

KURTOSIS AND SKEWNESS OF THE QUASI-RANGE DENSITY

n	r	α = 0		α = 1		α = 2	
		γ_1	γ_2	γ_1	γ_2	γ_1	γ_2
2	0	2.00000	6.00000	1.61985	3.79592	1.45143	2.93129
3	0	1.60997	4.08000	1.28467	2.57601	1.12746	1.94318
4	0	1.46356	3.48105	1.17289	2.25869	1.02405	1.71766
	1	2.00000	6.00000	1.66736	3.93619	1.54415	3.23786
5	0	1.38664	3.19367	1.12128	2.13106	0.97975	1.64239
	1	1.49342	3.44379	1.20124	2.12266	1.08368	1.64517
6	0	1.33922	3.02597	1.09335	2.06865	0.95817	1.61503
	1	1.28876	2.63316	1.02066	1.60233	0.90572	1.21081
	2	2.00000	6.00000	1.71662	4.16271	1.62036	3.58148
7	0	1.30706	2.91638	1.07672	2.03476	0.94705	1.60686
	1	1.17631	2.24305	0.92690	1.37556	0.81466	1.06160
	2	1.45600	3.23520	1.20108	2.06615	1.10796	1.67409
8	0	1.28382	2.83927	1.06617	2.01526	0.94134	1.60732
	1	1.10467	2.01580	0.87087	1.25444	0.76131	0.94163
	2	1.22889	2.34741	0.98982	1.44167	0.89710	1.12320
	3	2.00000	6.00000	1.75506	4.36257	1.67598	3.86413
9	0	1.26624	2.78211	1.05918	2.00378	0.93861	1.61199
	1	1.05538	1.86778	0.83434	1.18183	0.72798	0.89121
	2	1.10110	1.91773	0.87482	1.16113	0.78104	0.88453
	3	1.43985	3.14458	1.21521	2.09471	1.13820	1.75881
10	0	1.25248	2.73806	1.05441	1.99709	0.93766	1.61873
	1	1.01891	1.76403	0.80963	1.13420	0.70447	0.86165
	2	1.01858	1.66689	0.80159	1.00888	0.71036	0.76114
	3	1.20174	2.21931	0.98776	1.40312	0.90628	1.12585
	4	2.00000	6.00000	1.78479	4.52855	1.71765	4.09185

TABLE III

QUANTILES OF THE QUASI-RANGE DENSITY

$\alpha = 0$

n	r	p					
		0.005	0.010	0.025	0.050	0.100	0.500
2	0	0.00500	0.01002	0.02532	0.05128	0.10529	0.69307
3	0	0.07301	0.10538	0.17216	0.25311	0.38013	1.22796
4	0	0.18740	0.24300	0.34585	0.45969	0.62398	1.57842
	1	0.00250	0.00501	0.01266	0.02564	0.05264	0.34654
5	0	0.30863	0.38007	0.50671	0.64030	0.82655	1.83811
	1	0.04213	0.06072	0.09908	0.14544	0.21791	0.69315
6	0	0.42479	0.50748	0.64990	0.79688	0.99684	2.04451
	1	0.11740	0.15207	0.21580	0.28596	0.38639	0.95262
	2	0.00167	0.00334	0.00844	0.01709	0.03510	0.23102
7	0	0.53375	0.62354	0.77813	0.93379	1.14347	2.21539
	1	0.20479	0.25108	0.33334	0.41944	0.53824	1.15894
	2	0.02979	0.04291	0.07000	0.10271	0.15381	0.48732
8	0	0.63195	0.72900	0.89250	1.05498	1.27183	2.36134
	1	0.29324	0.34843	0.44390	0.54159	0.67272	1.33003
	2	0.08629	0.11183	0.15855	0.20991	0.28325	0.69314
	3	0.00125	0.00251	0.00633	0.01282	0.02632	0.17327
9	0	0.72437	0.82553	0.99556	1.16351	1.36586	2.48897
	1	0.37824	0.44069	0.54690	0.65251	0.79281	1.47626
	2	0.15514	0.18998	0.25139	0.31653	0.40525	0.86414
	3	0.02307	0.03323	0.05421	0.07952	0.11905	0.37660
10	0	0.81127	0.91547	1.08982	1.26206	1.48857	2.60200
	1	0.45961	0.52729	0.64153	0.75352	0.90082	1.60383
	2	0.22698	0.26926	0.34294	0.41730	0.51714	1.01016
	3	0.06841	0.08868	0.12568	0.16633	0.22431	0.54716
	4	0.00100	0.00200	0.00506	0.01026	0.03106	0.13861

TABLE III

QUANTILES OF THE QUASI-RANGE DENSITY

$\alpha = 0$

n	r	P				
		0.900	0.950	0.975	0.990	0.995
2	0	2.30240	2.99567	3.68498	4.60437	5.29340
3	0	2.96973	3.67609	4.37217	5.29507	5.98538
4	0	3.36629	4.07731	4.77571	5.69972	6.39047
	1	1.15120	1.49784	1.84249	2.30219	2.64670
5	0	3.64978	4.36286	5.06246	5.98700	6.67798
	1	1.63025	1.99957	2.35957	2.83150	3.18203
6	0	3.86992	4.58282	5.28502	6.20091	6.90103
	1	1.94796	2.32575	2.69276	3.16979	3.52290
	2	0.76747	0.99856	1.22832	1.53478	1.76445
7	0	4.04991	4.76439	5.46700	6.39209	7.08331
	1	2.18698	2.57036	2.94078	3.42061	3.77512
	2	1.13774	1.39188	1.63814	1.95970	2.19767
8	0	4.20322	4.91802	5.62091	6.54613	7.23741
	1	2.37942	2.76620	3.13680	3.62038	3.97576
	2	1.39982	1.66400	1.91877	2.24750	2.48940
	3	0.57560	0.74892	0.92126	1.15113	1.32342
9	0	4.33661	5.05119	5.75431	6.67967	7.37111
	1	2.54087	2.92966	3.30377	3.78656	4.14252
	2	1.60463	1.87563	2.13497	2.46779	2.71177
	3	0.87661	1.07119	1.25935	1.50448	1.68548
10	0	4.45434	5.16867	5.87193	6.79731	7.48870
	1	2.67921	3.07020	3.44561	3.92986	4.28718
	2	1.77394	2.04899	2.31117	2.64607	2.89041
	3	1.09902	1.30390	1.50080	1.75395	1.93967
	4	0.46045	0.59908	0.73688	0.92057	1.05808

TABLE III

QUANTILES OF THE QUASI-RANGE DENSITY

$\alpha = 1$

n	r	p					
		0.005	0.010	0.025	0.050	0.100	0.500
2	0	0.01003	0.02000	0.05000	0.10004	0.20122	1.14619
3	0	0.15027	0.21318	0.33959	0.48555	0.70285	1.96018
4	0	0.38050	0.48475	0.66685	0.85909	1.12332	2.47699
	1	0.00448	0.00896	0.02256	0.04554	0.09284	0.57346
5	0	0.61465	0.74254	0.95728	1.17211	1.45780	2.85266
	1	0.07813	0.11172	0.18037	0.26149	0.38491	1.12388
6	0	0.83269	0.97362	1.20669	1.43414	1.73171	3.14683
	1	0.22091	0.28193	0.39321	0.51212	0.67743	1.52622
	2	0.00287	0.00575	0.01450	0.02932	0.05996	0.37976
7	0	1.03410	1.17774	1.42216	1.65722	1.96187	3.38761
	1	0.38293	0.46442	0.60413	0.74625	0.93504	1.84107
	2	0.05287	0.07579	0.12284	0.17888	0.26488	0.79123
8	0	1.20573	1.36039	1.61054	1.85047	2.15960	3.59110
	1	0.54697	0.64110	0.80034	0.95622	1.15963	2.09924
	2	0.15572	0.19978	0.28006	0.36673	0.48740	1.11605
	3	0.00211	0.00423	0.01068	0.02160	0.04423	0.28334
9	0	1.36572	1.52221	1.77725	2.02043	2.33252	3.76710
	1	0.70230	0.80569	0.97815	1.14357	1.35663	2.31730
	2	0.28009	0.34052	0.44512	0.55203	0.69479	1.38255
	3	0.03998	0.05737	0.09316	0.13594	0.20188	0.61084
10	0	1.49910	1.66722	1.92535	2.17261	2.48593	3.92194
	1	0.84764	0.95677	1.13903	1.31120	1.53102	2.50571
	2	0.41212	0.48390	0.60595	0.72609	0.88332	1.60849
	3	0.12054	0.15496	0.21776	0.28582	0.38106	0.88207
	4	0.00167	0.00335	0.00845	0.01710	0.03503	0.22581

TABLE III

QUANTILES OF THE QUASI-RANGE DENSITY

$\alpha = 1$

n	r	P				
		0.900	0.950	0.975	0.990	0.995
2	0	3.27175	4.11295	4.92794	5.98947	6.77216
3	0	4.16697	5.00322	5.81462	6.86676	7.64312
4	0	4.69687	5.52970	6.33657	7.38306	8.15588
	1	1.72308	2.17995	2.62111	3.19314	3.61311
5	0	5.07275	5.90270	6.70622	7.74876	8.51912
	1	2.40290	2.87273	3.32185	3.89623	4.31575
6	0	5.36439	6.19088	6.99180	8.03133	8.79966
	1	2.85012	3.32395	3.77413	4.34775	4.76602
	2	1.17316	1.49210	1.80047	2.20036	2.49371
7	0	5.60113	6.42526	7.22407	8.26119	9.02826
	1	3.18586	3.66056	4.11055	4.68296	5.10014
	2	1.71395	2.05146	2.37266	2.78129	3.07633
8	0	5.80088	6.62250	7.41953	8.45465	9.22049
	1	3.45408	3.92910	4.37047	4.94955	5.36568
	2	2.09308	2.43834	2.76369	3.17538	3.47334
	3	0.86966	1.13587	1.37447	1.68435	1.91177
9	0	5.97180	6.79259	7.58813	8.62159	9.38646
	1	3.67748	4.15241	4.60109	5.17104	5.56644
	2	2.38878	2.73751	3.06497	3.47729	3.77511
	3	1.33711	1.60321	1.85610	2.17697	2.40930
10	0	6.12258	6.94204	7.73634	8.76652	9.53281
	1	3.86839	4.34278	4.79024	5.35765	5.77002
	2	2.63223	2.98384	3.31366	3.73052	4.03501
	3	1.66430	1.93972	2.19844	2.52363	2.75732
	4	0.71653	0.91741	1.11255	1.36651	1.55314

TABLE III

QUANTILES OF THE QUASI-RANGE DENSITY

$\alpha = 2$

n	r	p					
		0.005	0.010	0.025	0.050	0.100	0.500
2	0	0.01336	0.02667	0.06667	0.13337	0.26772	1.48193
3	0	0.20333	0.28564	0.45424	0.64757	0.93238	2.50734
4	0	0.50879	0.64998	0.89053	1.14226	1.48274	3.15140
	1	0.00585	0.01170	0.02945	0.05941	0.12098	0.73783
5	0	0.82962	0.99379	1.27530	1.55330	1.91741	3.61662
	1	0.10274	0.14676	0.23661	0.34240	0.50252	1.43883
6	0	1.11818	1.30041	1.60341	1.89481	2.27022	3.97875
	1	0.28977	0.37138	0.51652	0.67068	0.88390	1.94767
	2	0.00372	0.00745	0.01877	0.03794	0.07754	0.48703
7	0	1.37959	1.56939	1.88513	2.18369	2.56482	4.27413
	1	0.50378	0.61187	0.79431	0.97691	1.21834	2.34410
	2	0.06891	0.09870	0.15980	0.23238	0.34338	1.01193
8	0	0.02391	0.04771	0.11927	0.23854	0.47830	3.19375
	1	0.72115	0.84439	1.05052	1.25044	1.50886	2.66761
	2	0.20387	0.26082	0.36499	0.47662	0.63212	1.42440
	3	0.00272	0.00546	0.01376	0.02784	0.05696	0.36266
9	0	0.02525	0.05044	0.12609	0.25216	0.50498	3.37813
	1	0.01676	0.03332	0.08329	0.16657	0.33603	2.42899
	2	0.36761	0.44620	0.58122	0.71876	0.90124	1.76219
	3	0.05181	0.07431	0.12057	0.17576	0.26060	0.78040
10	0	0.02654	0.05303	0.13262	0.26519	0.53041	3.54876
	1	0.01752	0.03496	0.08741	0.17478	0.35075	2.59041
	2	0.01426	0.02803	0.07006	0.14066	0.29397	1.94924
	3	0.15658	0.20113	0.28240	0.37020	0.49250	1.12607
	4	0.00215	0.00430	0.01086	0.02197	0.04500	0.28867

TABLE III

QUANTILES OF THE QUASI-RANGE DENSITY

$\alpha = 2$

n	r	P				
		0.900	0.950	0.975	0.990	0.995
2	0	4.01037	4.96855	5.85349	7.05906	7.91742
3	0	5.07972	6.02017	6.92124	8.07669	8.92230
4	0	5.71182	6.64201	7.53318	8.67744	9.51617
	1	2.15116	2.69395	3.21071	3.87134	4.35110
5	0	6.15983	7.08213	7.96644	9.10309	9.93718
	1	2.98259	3.53200	4.05055	4.70579	5.17986
6	0	6.50574	7.42170	8.30084	9.43181	10.2625
	1	3.53788	4.07710	4.59285	5.24297	5.71319
	2	1.47396	1.86055	2.22963	2.70262	3.04588
7	0	6.78583	7.69744	8.57248	9.69897	10.5269
	1	3.93606	4.48300	4.99551	5.64059	6.10667
	2	2.14286	2.54586	2.92510	3.40241	3.74659
8	0	4.79530	5.01436	5.13140	5.20337	5.22781
	1	4.26202	4.80662	5.31629	5.95748	6.42078
	2	2.61016	3.01856	3.39975	3.87682	4.21984
	3	1.12121	1.42296	1.71239	2.08399	2.35403
9	0	4.97994	5.19647	5.31130	5.38197	5.40594
	1	3.60693	3.80386	3.91344	3.98470	4.00965
	2	2.97471	3.38592	3.76904	4.25049	4.60129
	3	1.67762	1.99891	2.30077	2.67848	2.94761
10	0	5.14947	5.36386	5.47697	5.54731	5.57095
	1	3.74602	3.92659	4.02941	4.09351	4.11570
	2	3.08666	3.38293	3.59500	3.76396	3.83384
	3	2.08667	2.41903	2.73064	3.12638	3.42082
	4	0.90482	1.15297	1.39213	1.70096	1.92703

TABLE IV

COMPARISONS WITH GHOSAL'S DATA

Data from Ghosal (Exponential)

	n = 5				n = 10			
r	μ_1'	μ_2	γ_1	γ_2	μ_1'	μ_2	γ_1	γ_2
0	2.0633	1.4236	1.3831	3.1940	2.8290	1.5398	1.2525	2.7723
1	0.8333	0.3611	1.4931	3.4210	1.7179	0.5274	1.0189	1.8508
2	-	-	-	-	1.0929	0.2618	1.0193	2.0234

Data from this paper ($\alpha = 0$)

	n = 5				n = 10			
r	μ_1'	μ_2	γ_1	γ_2	μ_1'	μ_2	γ_1	γ_2
0	2.0633	1.4236	1.3866	3.1937	2.8290	1.5398	1.2525	2.7381
1	0.8333	0.3611	1.4934	3.4438	1.7179	0.5274	1.0189	1.7640
2	-	-	-	-	1.0929	0.2618	1.0186	1.6669

TABLE V

COMPARISONS WITH GUPTA'S DATA

Absolute Deviations

n	r	α = 0	α = 1	α = 2
2	0	-	-	-
3	0	-	-	-
4	0	-	0.00001	0.00001
	1	-	-	-
5	0	-	-	-
	1	-	0.00001	-
6	0	-	-	-
	1	-	0.00001	0.00001
	2	-	-	0.00001
7	0	-	-	0.00001
	1	-	-	0.00001
	2	-	-	-
8	0	-	0.00001	0.00002
	1	-	-	0.00003
	2	-	0.00002	0.00002
	3	-	0.00001	0.00003
9	0	-	0.00001	0.00002
	1	-	0.00002	0.00003
	2	-	0.00004	0.00006
	3	-	0.00001	0.00004
10	0	-	0.00001	0.00004
	1	0.00001	0.00003	0.00018
	2	0.00002	0.00001	0.00023
	3	-	0.00007	0.00079
	4	-	0.00012	0.00028

TABLE VI

TESTS OF ACCURACY IN HIGHER MOMENTS

$\mu_i(n,r,\alpha)$

	i	Primary Program	Alternate Program
n = 10 α = 1 r = 2	1	1.69910	1.69922
	2	3.37648	3.37658
	3	7.67506	7.67516
	4	19.6406	19.6407
n = 10 α = 2 r = 2	1	2.14494	2.14516
	2	5.31528	5.31542
	3	14.8951	14.8952
	4	46.4916	46.4915

Appendix B

ON ESTIMATING THE SCALE PARAMETER OF THE RAYLEIGH DISTRIBUTION FROM DOUBLY CENSORED SAMPLES

Introduction

This paper is concerned with estimating the scale parameter of the Rayleigh distribution from censored samples. The Rayleigh distribution arises as a consequence of finding the resultant amplitude of several coplanar random amplitude vectors which are normally distributed (Siddiqui [1962]). Therefore it is useful in the analysis of acoustic data or other data obtained from measurements of amplitudes of electromagnetic waves received through a scattering medium. This distribution is also useful in communication engineering. Since for many reasons the samples could be censored, it is important to consider the analysis of such data.

The Rayleigh distribution is characterized by the probability density function (p.d.f.)

$$f(x) = \begin{cases} (2x/K)\exp\{-x^2/K\}, & 0 < x < \infty \\ 0 \quad , & \text{elsewhere} \end{cases} \tag{1.1}$$

for positive values of K, with expectation $\sqrt{K\pi}/2$ and variance $K(1-\pi/4)$. To estimate the parameter K, we employ the methods used by Tiku [1967, 1967(a), 1968, 1968(a), 1968(b)] for the censored samples from normal, exponential, logistic, log-normal distributions and progressively censored samples from normal distribution respectively, as described below.

The p.d.f. given in (1.1) can be reduced to

$$f_Z(z) = \begin{cases} 2z\exp(-z^2) \quad , & 0 < z < \infty \\ 0 \quad , & \text{elsewhere,} \end{cases} \tag{1.2}$$

where $x = Z/\sqrt{K}$

Let $g(z) = f_Z(z)/F_Z(z)$, where $F_Z(z)$ is the probability integral of $f_Z(z)$ given in (1.2). Then, over a small interval, say $a \leq x \leq b$, the linear approximation $\alpha + \beta z$ to $g(z)$ is a reasonable one, where α and β are constants such that

$$\beta = \{g(b) - g(a)\}/(b-a)$$

$$\alpha = g(a) - a\beta \quad .$$

Some numerical comparison between $g(z)$ and $\alpha + \beta z$ for various sample sizes are given in the example. The substitution $\alpha + \beta z$ for $g(z)$ in the likelihood equation results in a solution (K_C) which is easy to compute and which is asymptotically equivalent to the maximum likelihood estimator (MLE) according to Tiku's results [1967, 1967(a), 1968, 1968(a), 1968(b)]. To obtain greater accuracy, i.e., to get an approximation which is practically same as an actual MLE, it is suggested to try a linear approximation once again instead of using expensive and time consuming iteration procedures.

In the remainder of this paper the maximum likelihood equation for finding K_C will be set up and solved (for a doubley censored sample), expressions for the bias and variance of K_C will be developed, a numerical example will be given and an improved estimator of K_C (namely K_C') will be made.

Derivation of the Estimator K_C

Let $X_1, X_2, \ldots, X_n$ be a random sample from the Rayleigh distribution with the smallest $r_1 = q_1 \cdot n$ and the largest $r_2 = q_2 n$ observations being censored, where q_1 and q_2 are fixed and decided in advance (Type II censoring). The remaining sample values, arranged in order of magnitude, are

$$Y_{r_1+1}, Y_{r_1+2}, \ldots, Y_{n-r_2-1}, Y_{n-r_2},$$

i.e. forming a doubly censored Rayleigh sample of size $n - r_1 - r_2$. The probability density function of this censored sample (see, for example Saw [1961]) is

$$\frac{n!}{r_1! r_2!}(2/\sqrt{K})^{n-r_1-r_2} \left[\prod_{i=r_1+1}^{n-r_2} z_i\right] \exp\left\{-\sum_{i=r_1+1}^{n-r_2} z_i^2\right\} \cdot \left\{F_Z(z_{r_1+1})\right\}^{r_1} \left\{1-F_Z(z_{n-r_2})\right\}^{r_2} \quad (2.1)$$

where $z_i = y_i/\sqrt{K}$.

Taking logarithms of (2.1) and denoting it by L yields:

$$L = \log\left\{\frac{n!}{r_1!r_2!}\right\} + (n-r_1-r_2)\log(2/\sqrt{K}) + \sum_{i=r_1+1}^{n-r_2} \log z_i - \sum_{i=r_1+1}^{n-r_2} z_i^2 + r_1\{\log F_Z(z_{r_1+1})\} - r_2 z_{n-r_2}^2. \quad (2.2)$$

Taking the partial derivative of L with respect to K and simplifying

$$\frac{\partial L}{\partial K} = \left[\frac{r_1+r_2-n}{K}\right] + \frac{1}{K}\sum_{i=r_1+1}^{n-r_2} z_i^2 - \left[\frac{r_1 z_{r_1+1}}{2K}\right]\frac{f_Z(z_{r_1+1})}{F_Z(z_{r_1+1})} + \frac{r_2 z_{n-r_2}^2}{K}. \quad (2.3)$$

Setting $\frac{\partial L}{\partial K}$ equal to zero in equation (2.3) and solving for K would give the ordinary maximum likelihood estimator. However, this is a complicated nonlinear equation because of the term

$$g_Z(z_{r_1+1}) \equiv f_Z(z_{r_1+1})/F_Z(z_{r_1+1}) = \frac{2z_{r_1+1}\exp(-z_{r_1+1}^2)}{1-\exp(-z_{r_1+1}^2)}, \quad (2.4)$$

which is implicit in K, thus precluding any exact solution to (2.3). Hence an iteration procedure seems to be the only solution to this problem, but it is not only time-consuming and expensive for computation but also sometimes difficult to converge as is the case with the Rayleigh distribution*. Hence instead of an iteration procedure, we are proposing that $g(z_{r_1+1})$ be replaced by a linear approximation

$$g_Z(z_{r_1+1}) \simeq \alpha + \beta z_{r_1+1}. \quad (2.5)$$

Consider now

---

*We experienced divergences even with the actual population parameter as an initial guess when censoring is a little bit heavy.

$$\frac{\partial L}{\partial K} \approx \frac{\partial L'}{\partial K} = \frac{r_1+r_2-n}{K} + \frac{1}{K} \sum_{i=r_1+1}^{n-r_2} z_i^2$$

$$- \left[\frac{r_1 z_{r_1+1}}{2K}\right]\left[\alpha + \beta z_{r_1+1}\right] + \frac{r_2 z_{n-r_2}^2}{K} \qquad (2.6)$$

In this equation α and β are such that

$$\beta = \{g_Z(h_2) - g_Z(h_1)\} / (h_2-h_1) \text{ and}$$

$$\alpha = g_Z(h_1) - h_1 \beta \qquad (2.7)$$

where g is given by (2.4) and the interval (h_1,h_2) is chosen in such a way that z_{r_1+1} is sufficiently close to h_1 or h_2. But a difficulty arises, because we don't know the exact point of z_{r_1+1} since it depends on the parameter K we are going to estimate. This difficulty can be eliminated, for sufficiently large $n-r_1-r_2$, by choosing h_1 and h_2 so that

$$F_Z(h_1) = q_1 - \sqrt{\frac{1}{n} q_1(1-q_1)}$$

$$F_Z(h_2) = q_1 + \sqrt{\frac{1}{n} q_1(1-q_1)} \; . \qquad (2.8)$$

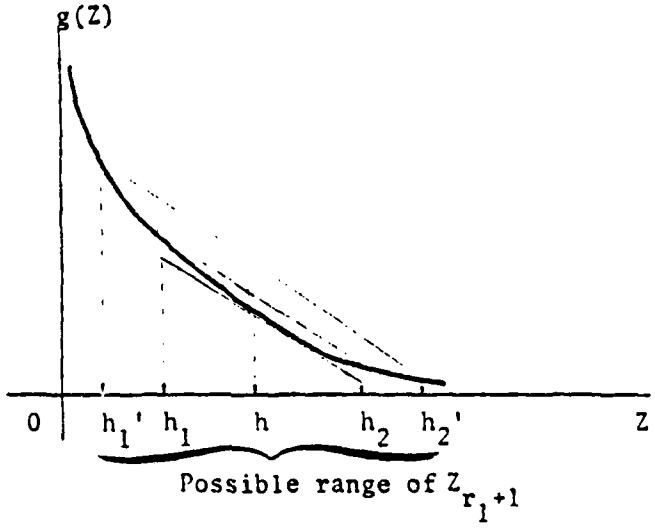
The reasoning behind this choice of interval end points is that it is logical to think of z_{r_1+1} as an estimate of the point below which $100 \cdot q_1$ percents of the population represented by $f_Z(z)$ lies. Also, the secant between h_1 and h_2 has smaller maximum error than the tangent at h or than the secant through h_1' and h_2' if one recalls that the probability is small that the Z_{r_1+1} will fall outside the interval h_1' and h_2', where $F_Z(h) = q_1$,

$$F_Z(h_1') = q_1 - 3\sqrt{\frac{1}{n} q_1(1-q_1)} \; , \qquad (2.9)$$

$$\text{and} \quad F_Z(h_2') = q_1 + 3\sqrt{\frac{1}{n} q_1(1-q_1)} \; .$$

(See figure 1).

fig. 1.



From (2.8) it can be seen that as n become large, $F_Z(h_1)$ approaches q_1 from the left at the same rate that $F_Z(h_2)$ approaches q_1 from the right. Thus the interval (h_1,h_2) shrinks to a single point, and α and β can be obtained by evaluating the derivative of $g_Z(z)$ at the point h, and in this case all 3 lines in the figure coincide. The degree of accuracy of the linear approximation is related to the width of the interval: The smaller the width of (h_1,h_2), the smaller the error of the approximation which is obvious from the figure. But one disadvantage is that the decrease of interval width h_2-h_1 is rather slow as n increases though the approximation was reasonably good when n is as small as 10 (see §4). Hence it is suggested in §5 that a possible acceleration of the approximation be used.

Equation (2.6) can now be solved analytically by setting $\frac{\partial L'}{\partial K} = 0$ and carrying out the algebra. After substitution of $y_i = \sqrt{K}\, z_i$ and some simplification, it follows that

$$\frac{1}{2K^2}(GK+B-D\sqrt{K}) = 0, \tag{2.10}$$

where $G = 2(r_1+r_2-n)$

$$B = 2\sum_{i=r_1+1}^{n-r_2} y_i^2 + 2\, r_2 y_{n-r_2}^2 - r_1 y_{r_1+1}^2 \beta$$

$$D = r_1 y_{r_1+1} \alpha.$$

In order to solve equation (2.10) for $K > 0$,

set $T = \sqrt{K}$ and rewrite it as a quadratic equation in T as

$$GT^2 - DT + B = 0. \tag{2.11}$$

Since $\sqrt{K}$ is positive we take the positive root of the equation (2.11) as T, i.e.

$$T = \frac{D - \sqrt{D^2-4GB}}{2G} .$$

Here note that G is negative and B and D are positive. Hence the estimator K_c is the square of T as

$$K_c = \left\{(D^2-2GB) - D\sqrt{D^2-4GB}\right\}/2G^2 \tag{2.12}$$

Properties of the Estimator K_c

The estimator K_c is same as an MLE except that a linear approximation was used to solve the likelihood equation. It is expected that the properties of K_c should be similar to those of an MLE. While calculation of the expected value of K_c from equation (2.12) would be extremely difficult, we can discuss the approximate conditional bias of K_c in the asymtotic case. Following Tiku [1967, p. 160], we have the approximate conditional bias given by

$$B_1 = E(\frac{\partial L'}{\partial K})/R^2(K)$$

where (3.1)

$$R^2(K) = -E(\frac{\partial^2 L'}{\partial K^2})$$

for large values of $n - r_1 - r_2$.

The bias B_1 can be calculated from the following equation:

$$E(\frac{\partial L'}{\partial K}) = \left[\frac{r_1+r_2-n}{K}\right] + (1/K) \sum_{i=r_1+1}^{n-r_2} E(Z_i^2) - (r_1\alpha/2K)\, E(Z_{r_1+1})$$

$$- (r_1\beta/2K)E(Z^2_{r_1+1}) + (r_2/K)\, E(Z^2_{n-r_2}). \tag{3.2}$$

Expected values of the order statistics from the Rayleigh distribution are given by (see Sarhan and Greenberg [1962])

$$E(Z_i) = \frac{n!}{(i-1)!(n-i)!} \sum_{J=0}^{i-1} \binom{i-1}{J} (-1)^J \left[\frac{\sqrt{\pi}}{2(n-i+1+J)^{3/2}} \right]$$

and

$$E(Z_i^2) = \frac{n!}{(i-1)!(n-1)!} \sum_{J=0}^{i-1} \binom{i-1}{J} (-1)^J \left[\frac{1}{(n-i+1+J)^2} \right]. \tag{3.3}$$

Differentiating (2.6) with respect to K and taking expectations and inserting the minus sign gives

$$\begin{aligned} R^2(K) &= -E\left(\frac{\partial^2 K}{\partial^2 K}\right) \\ &= \frac{r_1+r_2-n}{K^2} + \frac{2}{K^2} \sum_{i=r_1+1}^{n-r_2} E(Z_i^2) - \frac{3r_1\alpha}{4K^2} E(Z_{r_1+1}) \\ &\quad - \frac{r_1\beta}{K^2} E(Z^2_{r_1+1}) + \frac{2r_2}{K^2} E\left(Z^2_{n-r_2}\right) \end{aligned} \tag{3.4}$$

where expectations are also given by (3.3).

Also the asymptotic variance of K_c can be obtained from (3.4) by use of the asymptotic property (See Kendall and Stuart [1961]):

$$\mathrm{Var}(K_c) = \frac{1}{R^2(K)} \tag{3.5}$$

As Tiku [1967] justified, since K_c is an asymptotic MLE, its asymptotic properties, in no doubt, are the same as MLE. We can show that numerically, though they are complex, the bias B_1 in (3.1) would be zero when n is large. Furthermore, the attractiveness of the estimator is enhanced by the fact that it can be computed easily without worrying about divergence, without having to resort to iterations or procedures requiring expectations of order statistics (see Lloyd [1952]). Table 1 below shows the values of B_1 and $\mathrm{Var}(K_c)$ for various censoring when the sample size is 10 and K = 3.

Table 1. Asymptotic bias and variance of the estimator K_c for n = 10 with proportion q_1 censored from below and q_2 above.

q_1	q_2	(1)	(2)
.1	.2	0.019	0.469
.1	.4	- .092	.160
.1	.6	- .087	.203
.1	.7	- .055	.262
.2	.2	- .009	.687
.2	.4	- .167	.265
.2	.6	- .173	.363
.3	.4	- .217	.279
.3	.5	- .242	.403
.4	.4	- .055	.170

(1) = Bias/σ

(2) = $Var(K_c)/\sigma^2$, where

σ^2 = Variance of the Rayleigh distr. = $K[1-\pi/4)$

Numerical Example

Using the closed-form cumulative distribution function F(x) for the Rayleigh distribution and the probability integral transformation, a random sample of 100 observations was generated from the Rayleigh distribution with the population parameter K = 3.000. The sample of size 10 and 30 were chosen from the original sample of size 100.

The following analysis of data is given.

4.1 Linear approximation $\alpha + \beta z_{r_1+1}$ to $g(z_{r_1+1})$.

The following table shows the linear approximation $\alpha+\beta z_{r_1+1}$ to $g(z_{r_1+1})$ for various sample sizes with proportions q_1 and q_2 censored from the left and the right, respectively.

Table 2. n = 10

q_1	q_2	h_1	h_2	$g(y_{r_1+1}/\sqrt{K_e})$	$\alpha + \beta(y_{r_1+1}/\sqrt{K_e})$
.1	.1	.071	.465	7.09	15.65
.1	.2			6.33	13.82
.1	.3			6.33	13.82
.2	.1	.276	.628	4.16	4.98
.2	.2			3.67	4.39
.2	.3			3.67	4.39
.3	.1	.410	.767	4.16	4.28
.3	.2			3.67	3.94
.3	.3			3.67	3.94
.4	.1	.530	.899	3.66	3.49
.4	.2			3.19	3.22

Here, K_e is an actual MLE by iteration.

Table 3. n = 30

q_1	q_2	h_1	h_2	$g(y_{r_1+1}/\sqrt{K_e})$	$\alpha + \beta(y_{r_1+1})/\sqrt{K_e}$
.1	.1	.215	.410	8.65	8.84
.1	.2			7.78	8.27
.1	.3			7.20	7.83
.2	.1	.368	.564	4.56	4.69
.2	.2			4.04	4.24
.2	.3			3.71	3.90
.3	.1	.494	.695	4.06	3.89
.3	.2			3.58	3.58
.3	.3			3.27	3.34
.4	.1	.609	.820	2.83	2.79

Table 4. n = 100

q_1	q_2	h_1	h_2	g(z)	$\alpha + \beta z$
.1	.1	.269	.373	7.52	7.41
.1	.2			7.10	7.11
.1	.3			7.22	7.20
.2	.1	.417	.523	4.13	4.17
.2	.2			3.87	3.92
.2	.3			3.94	3.99
.3	.1	.541	.651	2.95	2.97
.3	.2			2.74	2.76
.3	.3			2.80	2.83
.4	.1	.657	.772	2.20	2.21
.4	.2			2.02	2.03

It can be seen from our earlier discussion, the approximation is better when n is large or h_2-h_1 is small but when the sample size is as small as 10, the approximation is not good; hence an improvement is suggested in §5.

4.2 Comparison With an actual MLE by iteration.

Table 5 shows comparison between K_c and an MLE by iterations* for n = 30.

Table 5.

q_1	q_2	K_c	MLE	Error
.1	.1	2.934	2.941	.002
.1	.2	2.380	2.399	.007
.1	.3	2.054	2.078	.011
.1	.4	1.887	1.917	.015
.1	.5	1.915	1.951	.018
.2	.1	2.938	2.956	.011
.2	.1	2.390	2.417	.012
.2	.3	2.071	2.097	.013
.2	.4	1.914	1.940	.016
.2	.5	1.946	1.978	.015
.3	.1	2.985	2.951	.011
.3	.2	2.411	2.410	.000
.3	.3	2.076	2.090	.006
.3	.4	1.909	1.931	.011
.3	.5	1.942	1.967	.012
.4	.1	3.003	2.990	.004
.4	.2	2.432	2.452	.007
.4	.3	2.105	2.132	.012
.4	.4	1.951	1.981	.015
.4	.5	1.991	2.029	.018

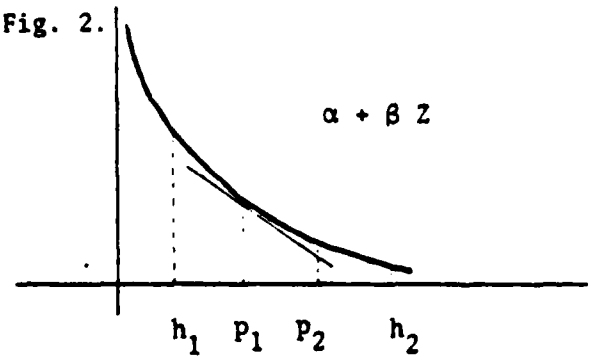
Error = $|K_c\text{-MLE}|/\text{MLE}$

*Iterated approximation is obtained by Newton-Raphson's method (See, for example, Dahlquist [1]).

Improvement by Linear Approximation Twice

Sometimes a greater accuracy for an approximation of an MLE is needed and as it was seen in §4, our estimator K_c is not sufficient for that purpose when the sample size is smaller than 30. To achieve greater accuracy, one solution is to use K_c as an initial guess and try an iteration procedure, which will give faster convergence in every case. But considering the computation time involved in iteration, it is suggested to try one more linear approximation using K_c as an estimator of K, which resulted in an estimator K_c' and showed almost the same accuracy (agreed with an actual MLE by iteration to two decimal digits when n = 10) as an MLE by iteration. The second linear approximation procedure is as follows: Since an estimator K_c of an actual MLE K_e is available, it is reasonable to think $y_{r_1+1}/\sqrt{K_c}$ lies closer to $y_{r_1+1}/\sqrt{K_e}$. Hence a tangent line at $y_{r_1+1}/\sqrt{K_c}$ can be used as an approximation of g(z) in the neighborhood of $y_{r_1+1}/\sqrt{K_c}$ (see Fig. 2).

Fig. 2.



$$p_1 = y_{r_1+1}/\sqrt{K_c}$$

$$p_2 = y_{r_1+1}/\sqrt{K_e}$$

i.e. new $\beta = g'(y_{r_1+1}/\sqrt{K_c})$ (5.1)

and $\alpha = g(y_{r_1+1}/\sqrt{K_c}) - (y_{r_1+1}/\sqrt{K_c}) \cdot \beta$,

and computing K_c by the same formula (2.13) as before will give a new estimator K_c'.

The following tables show the K_c' and MLE by iteration for sample size $n = 10$ and $n = 30$, $K = 3$.

Table 6. $n = 10$

q_1	q_2	K_c'	MLE	Error
.1	.1	2.5133	2.5123	.0003
.1	.2	2.0379	2.0369	.0004
.1	.3	.8669	.8669	.0000
.1	.4	.9576	.9575	.0000
.1	.5	.9576	.9575	.0000
.2	.1	2.5337	2.5335	.0000
.2	.2	2.0605	2.0603	.0001
.2	.3	2.0605	2.0603	.0001
.2	.4	.8947	.8944	.0004
.2	.5	.9915	.9912	.0002
.3	.1	2.5335	2.5335	.0000
.3	.2	2.0603	2.0603	.0000
.3	.3	2.0603	2.0603	.0000
.3	.4	.8943	.8943	.0000
.3	.5	.9912	.9912	.0000
.4	.1	2.4952	2.4952	.0000
.4	.2	2.0163	2.0163	.0000
.4	.3	2.0163	2.0163	.0000
.4	.4	.8265	.8264	.0001
.4	.5	.9117	.9113	.0004

Error = $|K_c'-MLE|/MLE$

Table 7. n = 30, K = 3.

q_1	q_2	K'_c	MLE	Error
.1	.1	2.94108	2.94108	.0
.1	.2	2.39997	2.39997	.0
.1	.3	2.07875	2.07874	.0
.1	.4	1.91790	1.91788	.1
.1	.5	1.95178	1.95175	.2
.2	.1	2.95687	2.95687	.0
.2	.2	2.41740	2.41738	.1
.2	.3	2.09744	2.09742	.1
.2	.4	1.94050	1.94047	.2
.2	.5	1.97899	1.97894	.3
.3	.1	2.95151	2.95148	.1
.3	.2	2.41096	2.41096	.0
.3	.3	2.09010	2.09010	.0
.3	.4	1.93130	1.93127	.1
.3	.5	1.96801	1.96796	.2
.4	.1	2.99032	2.99031	.0
.4	.2	2.45209	2.45207	.1
.4	.3	2.13266	2.13261	.2
.4	.4	1.98169	1.98160	.4
.4	.5	2.02922	2.02906	.8

Error = $(|K'_c - \text{MLE}|/\text{MLE}) \cdot (10^4)$

Table 8 and 9 show K_c' and MLE by iteration when K = 10.000.

Table 8. n = 10, K = 10.0

q_1	q_2	K_c'	MLE	Error
.1	.1	8.37781	8.37461	.00038
.1	.2	6.79320	6.78998	.00047
.1	.3	6.79320	6.78998	.00047
.1	.4	2.88993	2.88993	.00000
.1	.5	3.19218	3.19195	.00007
.2	.1	8.44567	8.44496	.00009
.2	.2	6.86851	6.86776	.00011
.2	.3	6.86851	6.86776	.00011
.2	.4	2.98254	2.98119	.00045
.2	.5	3.30493	3.30404	.00027
.3	.1	8.44497	8.44496	.00000
.3	.2	6.86789	6.86776	.00002
.3	.3	6.86789	6.86776	.00002
.3	.4	2.98122	2.98119	.00001
.3	.5	3.30424	3.30404	.00006
.4	.1	8.31778	8.31747	.00005
.4	.2	6.72130	6.72129	.00000
.4	.3	6.72130	6.72129	.00000
.4	.4	2.75501	2.75471	.00011
.4	.5	3.03910	3.03787	.00041

Error = $|K_c'-MLE|/MLE$

Table 9. n = 30, K = 10.000

q_1	q_2	K'_c	MLE	Error
.1	.1	9.80361	9.80361	.00000
.1	.2	7.99990	7.99989	.00000
.1	.3	6.92917	6.92914	.00000
.1	.4	6.39299	6.39293	.00001
.1	.5	6.50595	6.50583	.00002
.2	.1	9.85625	9.85623	.00000
.2	.2	8.05799	8.05793	.00001
.2	.3	6.99148	6.99141	.00001
.2	.4	6.46834	6.46824	.00002
.2	.5	6.59664	6.59646	.00003
.3	.1	9.83837	9.83827	.00001
.3	.2	8.03654	8.03654	.00000
.3	.3	6.96701	6.96699	.00000
.3	.4	6.43766	6.43756	.00001
.3	.5	6.56002	6.55988	.00002
.4	.1	9.96773	9.96770	.00000
.4	.2	8.17363	8.17356	.00001
.4	.3	7.10888	7.10871	.00002
.4	.4	6.60563	6.60535	.00004
.4	.5	6.76406	6.76354	.00008

Error = $(|K'_c - MLE|/MLE)$

Conclusion: The approximation K_c' to an MLE which was proposed in this section has some desirable advantages over the actual iteration procedure Considering the fact that the iteration procedure requires several steps, K_c' is clearly time saving because the procedure suggested requires the time equivalent to that which is needed for a single iteration.

The iteration procedure depends heavily on the initial guess and as such it is the main reason we do not want to compare the computing times required by our method and the iteration procedure. If the initial guess is poor the computing time required by the iteration will be lengthened. In fact, we experienced several divergence of the Newton-Raphson's method for our simple example though the initial guess was chosen as 2.0 when the actual parameter was 3. K_c' is computed using the previously obtained K_c and hence it has an advantage.

As regards the accuracy of the procedure suggested, we tried some more simulated data obtained from the distribution with different parameters and concluded that the relative error to the iterated estimate is negligible in most cases. (Less than 10^{-4} for the sample of size 10 or more). Hence we recommend to use K_c' when the time saving is a crucial matter or when there is a difficulty of setting the initial guess.

At present we are investigating the possibility of extending this new procedure to the results given in the series of papers by Tiku.

REFERENCES

Agrawal [72]: V. D. Agrawal and P. Agrawal, "An Automatic Test Generation for ILLIAC IV Logic Boards", IEEE Trans. on Comp., Sept.

Agrawal [75a]: P. Agrawal and V. D. Agrawal, "On Improving the Efficiency of Monte Carlo Test Generation," 1975 Internation Symposium on Fault-Tolerant Computing, June.

Agrawal [75b]: P. Agrawal and V. D. Agrawal, "Probabilistic Analysis of Random Test Generation Method for Irredundant Combinational Logic Networks", IEEE Trans. on Comp., July.

Armstrong [66]: D. B. Armstrong, "On Finding a Nearly Minimal Set of Fault Detection Tests for Combinational Logic Nets", IEEE Trans. of Electronic Comp., Feb.

Basu [64]: A. P. Basu, "Estimates of Reliability for some Distributions Useful in Life Testing", Technometrics, Vol. 6, pp. 215-19.

Benson [49]: R. Benson, "A Note on the Estimation of Mean and Deviation from Quantiles", J. Royal Stat. Soc., Ser. B, Vol II, pp. 91-100.

Bouricious [71]: W. G. Bouricious, W. C. Carter, D. Jessep, P. R. Schneider and A. B. Wadia, "Reliability Modelling of Fault-Tolerant Computers", IEEE Trans. of Comp., Nov.

Breuen [76]: M. A. Breuen and A. D. Friedman, Diopnosis and Reliable Design of Digital Systems, Computer Science Press.

Caldwell [53]: J. H. Caldwell, "The Distribution of Quasi-Ranges in Samples from a Normal Population", Ann. Math. Stat., Vol. 24, pp. 603-13.

Case [76]: G. R. Case, "A Statistical Method for Test Sequence Evaluation", The Proc. of 1976 Design Automation Conference.

Chu [57]: J. T. Chu, "Some Uses of Quasi-Ranges", Ann. Math. Stat., Vol. 28, pp. 173-80.

Cox [53]: D. R. Cox, "Some Simple Approximate Tests for Poisson Variates", Biometrika, Vol. 40, pp. 354-60.

Dahlquist [74]: G. Dahlquist, and A. Björk, Numerical Methods, Prentice-Hall, Inc., New Jersey, 218-251.

Drenick [60]: R. F. Drenick, "The Failure Law of Complex Equipment", J. Soc. Indust. Appl. Math, Vol. 8, No. 4, pp. 680-90.

Epstein [53]: B. Epstein, and M. Sobel, "Life Testing", JASA, Vol. 48, pp 486-502.

Epstein [54]: B. Epstein and M. Sobel, "Some Theorems Relevant to Life Testing From an Exponential Distribution", AMS, Vol. 25, pp. 375-81.

Fike [72]: John L. Fike, "Heunistic and Adaptive Techniques for Diagnostic Test Generation", Ph.D. Thesis, S.M.U.

Friedman [71]: A. D. Friedman and P. Menon, Fault Detection in Digital Circuits, Prentice-Hall, Inc.

Chosal [57]: A. Ghosal, "The Distribution of Quasi-Ranges in Samples from Rectangular and Exponential Distributions", J. Actuarial Student's Institute, Vol. 14, pp. 94-101.

Gupta [60]: S. S. Gupta, "Order Statistics from the Gamma Distribution", Technometrics, Vol. 2, pp. 243-62.

Harter [59]: H. L. Harter, "The Use of Sample Quasi-Ranges in Estimating Population Standard Deviation", Ann. Math. Stat., Vol. 30, pp. 980-99.

Kasouf [78]: G. Kasouf and S. Mercurio, "Evaluation of LSI/MSI Reliability Models", Proc. of 1978 Annual Reliability and Maintainability Symposium.

Kendall [61]: M. G. Kendall and A. Stuart, The Advanced Theory of Statistics Vol. 2, Hafner Publishing Co., New York, 42-44.

Lloyd [52]: E. H. Lloyd, "Least Squares Estimation of Location and Scale Parameters Using Order Statistics", Biometrika 39, 88-95.

Mann [75]: N. R. Mann, R. E. Schafer and N. D. Singpurwala, Methods for Statistical Analysis of Reliability and Life Data.

Mathur [71]: F. P. Mathur, "On Reliability Modelling and Analysis of Ultra Reliable Fault-Tolerant Digital Systems", IEEE Trans. on Comp., Nov.

Moreno [72]: V. Moreno, "A Logic Test Generation System Using a Parallel Simulaiton", Dept. of Comp. Sci., Univ. of Illinois, Urbana, ILLIAC Document 243.

Mosteller [46]: Frederick Mosteller, "On Some Useful Inefficient Statistics", Ann. Math. Stat., Vol. 17, pp. 377-408.

Parker [75a]: K. P. Parker, "Adaptive Random Test Generation", Tech note No. 73, Digital Systems Lab., Stanford Univ. Oct.

Parker [75b]: K. P. Parker, and E. J. McCluskey, "Production Treatment of General Combinational Networks", IEEE Trans. on Comp., June.

Parker [75c]: K. P. Parker and E. J. McCluskey, "Analysis of Logic Circuits with Faults Using Input Signal Probabilities", IEEE Trans. on Comp., May.

Pearson [20]: K. Pearson, "On the Probable Errors of Frequency Constants, Part III", Biometrika, Vol. 12, pp. 113-32.

Prescott [74]: P. Prescott, "Variances and Covariances of Order Statistics from the Gamma Distribution", Biometrika, Vol 61, pp. 607-13.

Pugh [63]: E. L. Pugh, "The Best Estimate of Reliability in the Exponential Case", J. of O.R.S.A., Vol. 11, pp. 56-61.

Rider [59]: P. R. Rider, "Quasi-Ranges of Samples from an Exponential Populaiton", Ann. Math. Stat., Vol. 30, pp. 252-54.

Roth [66]: J. P. Roth, "Diagnosis of Automata Failures: A Calculus and a Method", IBM Journal of Research and Development, Vol. 10, July.

Sarhan [62]: A. E. Sarhan, and B. G. Greenberg, Editors, Contributions to Order Statistics, John Wiley and Sons, Inc. New York, 12-27.

Saw [61]: J. G. Saw, "Estimation of Normal Population Parameters Given a Type Censored Sample", Biometrika 48, 367-77.

Schnurmann [75]: H. D. Schnurmann, E. Lindbloom and R. C. Carpenter, "The Weighted Random Test Pattern Generator", IEEE Trans. on Comp., July.

Sellers [68]: F. F. Sellers, M. Y. Hsiao and L. W. Bearson, "Analysing Errors with the Bollean Difference", IEEE Trans. on Comp., July.

Shedlesky [77]: J. J. Shedlesky, "Random Testing: Practicality vs. Verified Effectiveness", The Seventh International Conference on Fault-Tolerant Computing, June.

Siddiqui [62]: M. M. Siddiqui, "Some Problems Connected with Rayleigh Distributions", Journal of Research of the National Bureau of Standards 660, 167-174.

Su [74]: Stephen Y. H. Su, "Logic Design and Recent Developments, Part 5: Fault Diagnosis in Digital Networks", Computer Design, Jan.

Tees [71]: W. Tees, "Predicting Failure Rates of Yield-Enhanced LSI", Computer Design, Feb.

Tiku [67]: M. L. Tiku, "Estimating the Mean and Standard Deviation from a Censored Normal Sample", Biometrika 54, 155-165.

Tiku [67a]: M. L. Tiku, "A Note on Estimating the Location and Scale Parameters of the Exponential Distribution from a Censored Sample", The Australian Journal of Statistics 9, 49-54.

Tiku [68]: M. L. Tiku, "Estimating the Parameters of Normal and Logistic Distributions from Censored Samples", The Australian Journal of Statistics 10, 64-74.

Tiku [68a]: M. L. Tiku, "Estimating the Parameters of Log-Normal Distribution from Censored Samples", Journal of American Statistical Association 63, 134-140.

Tiku [68b]: M. L. Tiku, "Estimating the Mean and Standard Deviation from Progressively Censored Normal Samples", Journal of Indian Argricultural Statistics 20, 20-25.

Vodovoz [75]: Erwin Vodovoz, "Testing Microprocessors is a Gamble", Electronic Products Magazine, Nov.

Watkins [70]: W. B. Watkins, "Introducing Companies to Automated Testing", Electronic Packaging and Production, Oct.