Improvement of Sequential Tests in IEC-61124 Reliability Testing – Compliance Tests for Constant Failure Rate

Ofer Shaham
Microelectronics Directorate
RAFAEL – Advanced Defense Systems Ltd
Haifa, Israel
Email: ofersh@rafael.co.il

Yefim Haim Michlin
Faculty of Industrial Engineering & Management
Technion - Israel Institute of Technology
Haifa, Israel
Email: yefim@technion.ac.il

***Abstract* — An improvement of IEC-61124 in the field of a sequential probability ratio test (SPRT) is proposed. The current standard does not provide a proper solution for modern industry's needs, and the test plans are not up-to-date with the knowledge in the area of sequential tests. The advantages of the proposed version are reflected by the efficacy and accuracy of the tests, the wider range of the ready to use test's parameters, and available data regarding the test's characteristics. The proposed version is a significant improvement over the existing one. The changes will extend the use of SPRT and this standard. The proposal for updating the standard has been accepted to the work-plan of TC-56 of IEC.**

Keywords — Exponential distribution; compliance; mass-production; reliability; sequential probability ratio test.

I. INTRODUCTION

The main purpose of the work is to improve the standard IEC 61124 [1] in the field of sequential tests [2]–[5]. Sequential tests have significant importance and actuality in quality assurance and reliability [6] [7]. Today's methods and computing tools permit better planning of sequential tests, in accordance with the latest practical demands (more stringent over time) [8]–[17].

The proposed version will include shorter tests (more economically efficient), more accurate tests, a wider range of test plans, and significant additional characteristic data. It will also address the complex and stringent needs of today's industry.

The sequential method of testing is described as follows. A rule is given for making one of the following three decisions at any stage of the experiment: (1) to accept (2) to reject (3) to continue the experiment by making an additional observation. The process is continued until either the first or the second decision is made.

An essential feature of the sequential test is that the required number of observations depends on their outcome and is not predetermined, but a random variable. Advantage of the sequential test lies in the fact that its expected accumulated test time to decision (ETT) is minimal at two typical points of its operating characteristic (OC) representing the error probabilities of the I- and II-types (α and β) [3] [4].

The characteristics of this test are obtainable from its boundaries [8] by means of Aroian's [18] well-known direct method, following an idea outlined earlier by Barnard [19].

II. MOTIVATION

Sequential tests have a significant importance and actuality in industrial acceptance sampling, in information technology and in reliability examination. The improvement is manifested in the means of the range, truncation and ETT, accuracy, user interface, usability, and in the simplicity of the planning. Accurate data for the test characteristics are part of both today's needs and ability.

A. Disadvantages of the current standard

The standard does not provide a proper solution for modern industry's needs, and the test plans are not up-to-date with the knowledge in the area of sequential tests.

- General:
  - Insufficient range of the test parameters (risks and discrimination ratio). Only a total of 17 "ready to use" sequential test plans are given.
  - The test's truncation is not optimal and considerations for truncation are not homogenous.
  - Lack of information for the relationship between the test truncation time and the expected test time.
  - The method for presentation of the tests (figures and detailed tables in Annex A and D) are suitable only for a few tests (not suitable for many). It limits the number of tests to be displayed in the standard.
  - For some of the test plans (A.4, A.7, A.9, C.8) in the figures of the "accept and reject" lines Y- axes shows non-integer values for the observed number of failures. This is an editorial error.
- Test plans A.1 – A.9:
  - The true producer's and consumer's risks are wrong and far from the nominal.
  - Non-optimal truncation - tests with lower maximum duration and lower expected time are available.
  - Tests are limited to three equal nominal risks only (10%, 20% and 30%).
  - Additional test plans can be calculated by formulas (given in Annex E). The formula leads to substantial deviation (unknown) from the nominal risks and does not carry out optimal truncation.

- Test plans C.1 – C.8:
  - Non-optimal truncation - the tests are overly truncated, and as a result, have high expected time, so that the sequential tests' advantages are lost.
  - The plans are limited to eight tests with *D*=1.7 only.

The standard does not provide the needs of modern industry, and the test plans are not up-to-date with the knowledge in the area of sequential tests.

B. Advantage of the proposed version

The proposed version will address the complex and stringent needs of today's industry. It will bring to the fore the latest knowledge and methods in truncated sequential tests.

1) The tests are substantially truncated (the maximum test duration is low) without significantly increasing the ETT [8].
2) The true producer's and consumer's risks are given and always very close to the nominal.
3) A wider range of the test parameters (risks and discrimination ratio) are given—a total of 60 tests.
4) The test plans include a series of unequal risks for producer and consumer.
5) Five accurate values of ETT are given in the test plan table for fulfilling all practical needs (easy to use with spreadsheet for full test characteristics presentation).
6) Suitable for interpolation by risks.

The proposed version addresses the complex and stringent needs of today's industry. It brings to the fore the latest knowledge and methods in truncated sequential tests.

III. DEMONSTRATION AND COMPARISON

This section deals with the examination of the standard's sequential test plans. The analysis included accurate calculation of the true characteristics of the tests A and C. In order to perform the comparison, corresponding test plans were calculated by using the advanced methodology [8]. The comparison focused on the differences in the parameter values (and accuracy), test characteristics, and available data. It expressed the significant advantage of the proposed tests in formulating the new version, as is demonstrated in the following subsections.

A. Optimal truncation of the tests

In the absence of direct relationships between the test's parameters and the test's expenses and efficacy, for given *D*, α, β, it is possible to consider the following as the main test's efficacy factors:

- expected test time function (*ETT*) as a multiple of m_0;
- accumulated test truncation time (T^*_t/m_0) (maximum test duration) as a multiple of m_0;
- test truncation failure number (r_0).

The truncation selection is composed of two considerations:

a) In general, for sequential tests, as the truncation is heavier (lower max duration), the ETT is higher.
b) Optimal ratio between the max test duration and the max failure number permits heavy truncation without increasing the ETT function [17] [8].

In tests A, the ratio between the max test duration and the max failure number is not the optimal.

In tests C, the tests have very strong truncation while ETT is substantially higher than the non-truncated (Wald's [3]) tests.

In the proposed tests, the truncation is the heaviest without a significant increase of ETT (vs. the non-truncated). See Figs. 1 and 2. The advantage of the tests over the current ones regarding test time is illustrated (as an example) in Figs. 1 and 2 (tests A.8 and C.1 correspondingly). In order to permit a comparison between the tests by their time to decision, corresponding tests with the same true risks were designed.

In Fig. 1, the test truncation failure number (r_0) of the alternative test is better (7 vs. 8 of the standard's A.8), the maximum duration is also better (1.5% shorter) and still, its ETT is better. The ETT of test A.8 is 12% higher at the average vs. the ideal non-truncated test, while the proposed is only 6%.

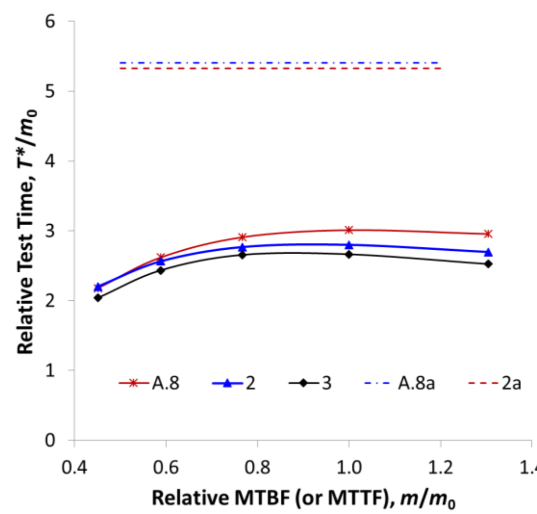

Figure 1. Expected test time and truncation time for A.8 and the alternative.
Test data: *D*=1.5, α=β=0.30 (nominal risks), α'=0.289 β'=0.363 (true risks for both tests).
A.8 – ETT (Expected accumulated test time to decision, T^*_e/m_0) of A.8;
2 – Ditto alternative;
3 – Ditto non-truncated test;
A.8a – Accumulated test truncation time (T^*_t/m_0) of A.8;
2a – Ditto alternative.

In Fig. 2, r_0 of the proposed test is 43 vs. 39 of the standard's C.1 (the maximum duration is slightly higher by 10%), but its ETT is much lower. The ETT of test C.1 is 41% higher at the average vs. the ideal non-truncated test, which doubts the relevance of the test. The proposed test is very close to the ideal non-truncated test (only 10% higher at the average).

Examining all C tests vs. the proposed shows a 20% reduction of the ETT (at the average).

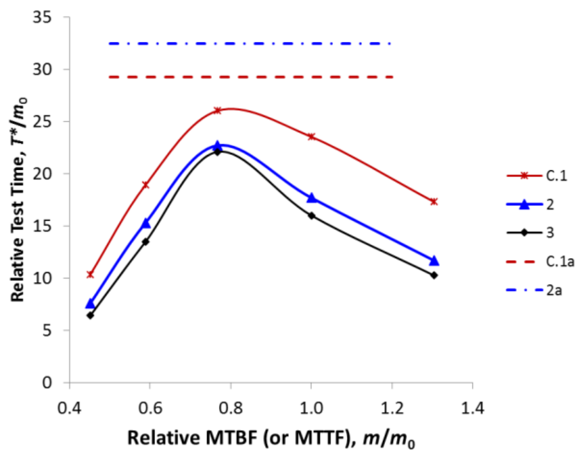

Figure 2. Expected test time and truncation time for C.1 and the proposed alternative.

Test data: D=1.7, α'=β'=0.050 (true risks).

C.1 – ETT (Expected accumulated test time to decision, T^*_e/m_0) of C.1, r_0=39;

2 – Ditto alternative, r_0=43;

3 – Ditto non-truncated test, r unlimited;

C.1a – Accumulated test truncation time (T^*_t/m_0) of C.1, r_0=39;

2a – Ditto alternative, r_0=43.

The advantage of the proposed tests (significantly truncated without significantly increasing the ETT vs. the non-truncated test) is achieved due to the optimal ratio between the test truncation failure number (r_0) and the accumulated test truncation time (T^*_t).

B. Range of the test parameters (variety of test plans, incl. interpolation)

The variety of "ready to use" tests in the current standard is very limited (17 total):

- Nine A tests that are restricted to equal nominal risks only (α=β).
- Eight C tests that are limited to D=1.7.

Additional A tests can be calculated by formulas (given in Annex E). The formulas result in substantial unknown deviations from the nominal risks and they do not carry out optimal truncation.

For additional C tests the standard refers to GOST R 27.402 [20].

- The GOST 27.402-95 includes 14 SPRT tests for D and α=β combinations.
- A method for additional test calculations is enclosed in Annex K of IEC-61124:
  - It consists of a complicated iterative procedure of finding values of four unknowns.
  - The end of the procedure is detected by an ambivalent variance of ETT function (change in the unknown values can lead to an increase of the ETT function on one side and a decrease on the other, so it is impossible to define an optimal test and its parameters).
  - No data is available regarding the optimal truncation (time and failure number) vs. the ETT function.

The proposed version includes 60 test plans that are "ready to use". It provides tests with a variety of risk values including several risk ratios (Table I), over few discrimination ratios: D = 1.5, 1.7, 2, 3, 5.

TABLE I. NOMINAL RISKS

α	0.05			0.1				0.2			0.3	0.4
β	0.05	0.1	0.2	0.05	0.1	0.2	0.4	0.05	0.1	0.2	0.3	0.1

Simple interpolation formulas for additional test plans are proposed. Using the formulas, it is possible to calculate test plans with intermediate parameter values with high accuracy (much more exact than in the standard's Annex E).

C. Risk accuracy

The sequential A test plans in the standard features two kinds of incorrectness regarding the test's risks:

- The true risks significantly deviate from the nominal (as seen in the standard's Table II)
- The "true" values declared in the standard's Table II are not accurate (the risks relative error is up to 104% when checked).

Detailed analysis of standard's 2nd edition (2006) is available in [17], and valid for the 3rd edition as well.

The standard's Type C plan features high accuracy of the risks.

In the proposed version, the true values are nominal (the average relative deviation is less than 0.002%).

D. Expected accumulated test time to decision (ETT)

- Type A plan's characteristics in the standard have substantial errors in the ETT (up to 17%) [17].
- In the current standard, the ETT is given by graphs (designated for each test plan). This approach is suitable only for a limited number of tests and is not necessary for practical use.
- The proposed test plans include five accurate values of ETT vs. m in the region between $m_1/D^{0.5}$ and $m_0*D^{0.5}$ (a constant step on a logarithmic scale). The given data is enough to restore the ETT function with high accuracy in this region. See Table II and Fig. 3.
- A spreadsheet (in accordance with Annex F of the standard) for generating the graphs and the other test characteristics (OC and boundaries) will be attached to the proposed version.

- For any of the proposed test plans (see Table III), the ETT curve is defined via five points vs. the true *m* (MTBF or MTTF), and according to Table II:

TABLE II. ETT VS. TRUE *m* (MTBF OR MTTF)

m	$m_1/\sqrt{D}$	$m_1 = m_0/D$	$m_0/\sqrt{D}$	m_0	$m_0\sqrt{D}$
ETTj	ETT_L	ETT_1	ETT_M	ETT_0	ETT_H

Note: j = (L, 1, M, 0, H)

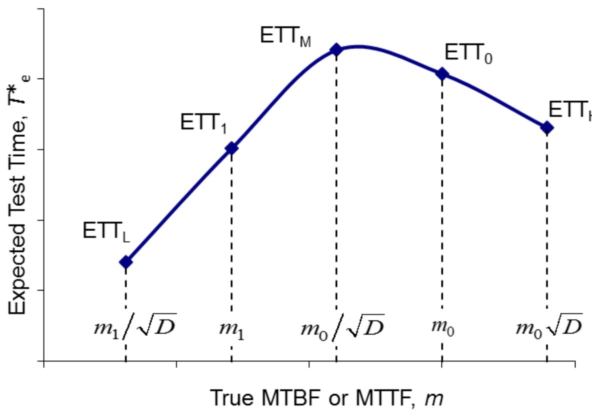

Figure 3. Example of the curve of expected test time to decision (ETT).

D. Operational characteristics (OC)

- In the current standard, the OC is given by graphs (designated for each test plan). This approach is suitable only for a limited number of tests and is not necessary for practical use.
- For the proposed test plans, the OC is in accordance to Wald's formulas as presented in Annex E.3.2 and Annex F.2.2 (for construction of the OC graph by spreadsheet program). The OC by Wald's formulas is accurate only for true risks (α' and β', see 4.3).

IV. PROPOSED TEST PLANS PRESENTATION

The proposed version includes 60 tests plans with all the required test data (boundaries and characteristics) in one table (see Table III).

The proposed changes can be implemented as follows:

- As an updated version of the standard in the part of sequential tests.
- Annex to the current standard, as an additional option for the sequential test plans.

V. CONCLUSION

The proposed version will be a significant improvement over the existing one. It is the result of development in the field of SPRT in recent years and the available computing power, which will support the requirements of today. It is both possible and necessary to conduct more economically efficient, more precise and more complex tests because of the accession of computer systems and stricter requirements in quality assurance and reliability. The changes will extend the use of SPRT and this standard.

The proposal for updating the standard has been accepted to the work-plan of Technical Committee TC-56 of International Electrotechnical Commission (IEC).

ACKNOWLEDGEMENT

The authors are indebted to Mrs. E. Leshchenko and to Mrs. M. Zenevich for assistance with the computer program and calculations.

The research project was supported by the Israel Ministry of Absorption and the Planning and Budgeting Committee of the Israel Council for Higher Education.

This paper is dedicated to the bright memory of Eng. Eliezer Goldberg (1918–2015) in honor of his work over many years of phenomenal editing.

REFERENCES

[1] IEC 61124:2012, *Reliability testing – Compliance test for constant failure rate and constant failure intensity*.

[2] IEC 61123:1991, *Reliability testing – Compliance test plans for success ratio*.

[3] A. Wald, *Sequential Analysis*. NY: John Wiley & Sons, pp. 221, 1947.

[4] A. Wald and J. Wolfowitz, "Optimum character of the sequential probability ratio test," *Ann. of Math. Stat*., vol. 19, no 3, pp. 326–339, 1948.

[5] B. Eisenberg and B. K. Ghosh, "The sequential probability ratio test," in *Handbook of Sequential Analysis*, B. K. Ghosh and P. K. Sen, Eds. NY: Marcel Dekker, pp. 47–66, 1991.

[6] K.C. Kapur and L.R. Lamberson, *Reliability in Engineering Design*. NY: Wiley, 1977.

[7] R. Kenett, Sh. Zacks, and D. Amberti, *Modern Industrial Statistics: With Applications in R, MINITAB and JMP*: John Wiley & Sons, Ltd., 2014.

[8] Y. H. Michlin and O. Shaham, "Planning of truncated sequential binomial test via relative efficiency," *Quality and Reliab. Engineering Int.* vol. 29: 369–383, 2013.

[9] Y. H. Michlin and G. Grabarnik, "Search boundaries of truncated discrete sequential test," *J. Appl. Stat.*, vol. 37, no. 5, pp.707–724, 2010.

[10] Y. H. Michlin, G. Grabarnik, and E. Leshchenko, "Comparison of the mean time between failures for two systems under short tests," *IEEE Trans. Reliab.*, vol. 58, no. 4, pp. 589–596, 2009.

[11] Y. H. Michlin and G. Grabarnik, "Sequential testing for comparison of the mean time between failures for two systems," *IEEE Trans. Reliab.*, vol. 56, no. 2, pp. 321–331, June 2007.

[12] Y. H. Michlin, D. Ingman, and Y. Dayan, “Sequential test for arbitrary ratio of mean times between failures,” *Int. J. of Operations Research and Information Systems*, vol. 2, no. 1, pp. 66–81, 2011.

[13] Y. H. Michlin, D. Ingman, and V. Kaplunov, “Sequential testing for two exponential distributions at arbitrary risks,” *Int. J. of Quality and Reliability Management*, vol. 29, no. 4, pp. 451–468, 2012.

[14] N. Mukhopadhyay and B. M. de Silva, *Sequential Methods and Their Applications*. Boca Raton FL: CRC Press, 2009.

[15] Y. H. Michlin and O. Shaham, “Remarks on the Paper "Closed Sequential and Multistage Inference on Binary Responses with or without Replacement" by Ignatova, Deutsch, and Edwards - The American Statistician, vol. 66, 2012” Comment by Michlin and Shaham and Reply, *The American Statistician*, vol. 68 no. 2, pp. 128-128, 2014.

[16] Y. H. Michlin, O. Shaham, and Ya. P. Lumelskii, “Substantiation of sequential test parameters for mass-produced electronic devices,” *2012 IEEE 27-th Convention of Electrical and Electronics Engineers in Israel*, Eilat, Israel, 2012.

[17] Y. H. Michlin, L. Meshkov, and I. Grunin, “Improvement of "Sequential testing" sections of MIL-HDBK-781A and IEC 61124,” *IEEE Trans. Reliab.*, vol. 57, no. 2, pp. 379–387, 2008.

[18] L. A. Aroian, “Sequential analysis – direct method,” *Technometrics*, vol. 10, pp. 125–132, 1968.

[19] G. A. Barnard, “Sequential test in industrial statistics,” *J. R. Statist. Soc.*, Suppl., 8, pp. 1–21, 1946.

[20] GOST R 27.402:95, *Dependability technics – Compliance test plans for mean operating time to failure or between failures – Part 1: Exponential case*.

ABBREVIATIONS

(The singular and plural of an abbreviation are always spelled the same)

ETT	expected accumulated test time to decision
MTBF	mean operating time between failures
MTTF	mean time to failure
OC	operating characteristic
SPRT	sequential probability ratio test

SYMBOLS

a	the accept line’s intersection with the r axis
b	the accept and reject line’s slope
b^*	$b^* = b \times m_0$, slope
c	the reject line’s intersection with the r axis
D	discrimination ratio; $D=m_0/m_1$
m	true mean operating time between failures (MTBF) or mean time to failure (MTBF)
m_0	specified MTTF or MTBF (design goal)
m_1	lower limit for MTTF or MTBF
T^*_e	expected accumulated test time to decision (ETT)
T^*_t	accumulated test time stated as termination criterion (truncation or max duration)
r	observed number of failures during the test
r_0	test truncation failure number
α	nominal producer’s risk (type I risk)
α’	true producer’s risk
β	nominal consumer’s risk (type II risk)
β’	true consumer’s risk

TABLE III. EXAMPLE FOR SEQUENTIAL TEST PLANS

D, b^*	α’ %	β’ %	a	c	T^*_t/m_0	r_0	ETT_L	ETT_1	ETT_M	ETT_0	ETT_H
D=1.5 b^*=1.2333	5	5	-8.070	7.905	63.93	78	13.6	27.0	40.8	30.5	19.3
	5	10	-6.465	8.054	48.37	60	13.8	24.8	31.8	23.7	15.4
	5	20	-4.455	7.506	36.22	46	12.5	19.5	21.5	15.8	10.5
	10	5	-8.137	6.395	48.69	58	11.1	21.3	32.1	27.9	19.2
	10	10	-6.209	6.054	37.96	46	10.4	18.3	24.2	20.6	14.5
	10	20	-4.476	5.960	24.97	31	9.8	14.5	16.4	13.8	10.3
	10	40	-2.455	4.593	13.70	18	6.3	7.6	7.7	6.6	5.3
	20	5	-7.516	4.296	37.39	43	7.6	14.2	21.8	21.9	17.0
	20	10	-5.811	4.361	25.92	30	7.5	12.4	16.4	16.0	12.8
	20	20	-3.980	3.921	16.09	19	6.3	8.8	10.2	9.8	8.3
	30	30	-2.320	2.354	6.44	7	3.0	3.5	3.9	3.9	3.7
	40	10	-4.902	2.260	14.26	15	4.0	6.0	7.9	8.8	8.6

 ISBN: 978-1-61208-593-7