

Credibility Evaluation of Operational Test Simulation Data under Small Sample Circumstance

Zhu Zhanfei^{1,*}, Han Xinwen¹, Li Wensheng², Yang Shutao¹, and Wang Bingchuan¹

¹Electronic Countermeasures Training, Engineer, Luo Yang of Henan, China

²Electronic Countermeasures Training, Senior Engineer, Luo Yang of Henan, China

Abstract. It is highly necessary to study how to analyze the reliability of simulation data under small sample circumstance when the number of times live operational test is strictly limited. Based on the analysis of existing test ideas and methods, combined with the characteristics of sequence statistics of uniform distribution, a new method of consistency verification is proposed by constructing the variable-scale differential quotient sequence statistic. The research shows that this method is not limited by the sample size, and the credibility of the simulation data can be quickly judged by MATLAB programming.

1 Introduction

With the constant development and perfection of modern modeling and simulation technology, it has gradually become an important class of methods for operational testing that the combat effectiveness and operational suitability test of weaponry and equipment are carried out on the distributed networked computer system acquiring the complete solution of the entire process by constructing a real-life weaponry and battlefield environment [1]. But the model is, after all, an approximation of the system, or even a simplification of the system. The resulting data must be validated by real test data before it can be used for subsequent analysis and evaluation. However, due to the constraints of the test period, environmental conditions and funding guarantee, the times of live operational tests actually carried out on some large-scale weapon systems is often strictly limited and the resulting data show obvious features of small sample. Therefore, it is highly necessary evaluating the credibility of operational simulation data under small sample circumstance.

The common idea of credibility verification is to compare the difference between simulation and live test results under the same input (test) conditions, and verify the consistency of simulation data and live data, which is called consistency test. According to the probability theory and mathematical statistics theory, consistency test can be understood as testing whether two sets of data obey the same distribution. The commonly used classical significance test methods, such as the Kolmogorov-Smirnov method and rank sum test, generally requiring adequate experimental data, are not suitable for use under small sample circumstance. This paper will improve the test ideas and methods based on the data consistency research results [2] [3] [4], and analyze the credibility of the simulation data in detail.

2 Existing test ideas and methods

Suppose $x_{1t}, x_{2t}, \dots, x_{Mt}$ is the data sample obtained through the simulation operation (M is generally relatively large), x_1, x_2, \dots, x_n is the data sample obtained by the live test (n is usually relatively small), and now it is required to test whether the two sets of data obey the same distribution. If the simulation data and the live test of small sample data mixed together, then tested, the results often turn out to be accepting assumptions, which cannot reflect the difference. Because many of simulation information overwhelmed the live test information, even if there are differences, it is difficult to reflect.

In Literature [2], $x_{1t}, x_{2t}, \dots, x_{Mt}$ is first divided into L groups randomly, where: $L = \lfloor M/n \rfloor$ when $M > n$; $L = 1$ when $M \leq n$. After that, the consistency between the simulation data and the live test data is verified one group by one group. Then the reliability of the simulation data under the small sample circumstance is transformed into the hypothesis test problem of whether the small sample data are subject to the same distribution. Let y_1, y_2, \dots, y_m be one of the L groups of data, and the 2 groups data of sub-samples x_1, x_2, \dots, x_n and y_1, y_2, \dots, y_m will be mixed and sorted from small to large to form a sequential sub-sample, denoted by $x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(N)}$, $N = n + m$. The probability density of $x_{(k)}$ can be proved to be:

* Corresponding author: 454086276@qq.com

$$\frac{d}{dx} P\{x_{(k)} \leq x\} = \frac{N!}{(k-1)!(N-k)!} [F(x)]^{k-1} [(1-F(x))]^{N-k} f(x)$$

$$, \quad k = 1, 2, \dots, N$$

Where: $F(x)$ is the distribution function, $f(x)$ is the distribution density. In the case that the distribution parameters of $F(x)$ are known, the sequence of random variables corresponding to the sequential sub-sample is denoted as α_k :

$$\alpha_k = \int_{-\infty}^{x_{(k)}} \frac{N!}{(k-1)!(N-k)!} [F(x)]^{k-1} (1-F(x))^{N-k} f(x) dx$$

$$= \int_0^{F(x_{(k)})} \frac{N!}{(k-1)!(N-k)!} t^{k-1} (1-t)^{N-k} dt \quad (1)$$

The random variable α_k can be seen as the sequence statistic of the sub-samples taken from the uniform distribution in the $[0, 1]$ interval. Therefore, the consistency problem between the two sub-samples x_1, x_2, \dots, x_n and y_1, y_2, \dots, y_m is transformed into checking the problem of whether the random sequence α_k is subject to $(0, 1)$ uniform distribution.

Literature [3] questioned the method of checking that the random sequence α_k obeys the uniform distribution of $(0,1)$ in Literature [2], and proved its irrationality by analysis, and then used the statistic

$$K_{ijk} = \frac{x_{(j)} - x_{(i)}}{x_{(k)} - x_{(i)}} \quad (1 \leq i < j < k \leq N)$$

to analyze the compatibility of the small sample capacity of normal distribution. The physical meaning of K_{ijk} lies in that it uses variable-scale measurements to extract large amounts of information from small samples, measures tiny differences between data, and amplifies these differences to identify different features between them. When the sample capacity $N = 3$, the distribution

$$\text{function } F(z) = \frac{3}{\pi} \left[\arctan\left(\frac{2z-1}{\sqrt{3}}\right) + \frac{\pi}{6} \right]$$

of the statistic K_{123} can be solved. Then for a given significance level α , α -quantile, the upper α -quantile or two-sided α -quantiles can be obtained according the rejection region of the testing. Finally the consistency between the data can be determined by comparing the values of K_{123} and the quantiles. However, when the sample capacity $N \geq 4$, the accurate analytic expression without the integral of the statistic K_{ijk} distribution cannot be obtained, and only the critical values of the test can be obtained through a large number of computer simulation operations, which greatly reduces the efficiency of the test, So it is necessary to improve the above method, giving a simple and feasible method of consistency test.

3 Improved test ideas and methods

It can be seen from the above that the irrationality of Literature [2] lies in the idea and method of checking that α_k is subject to the uniform distribution of $(0, 1)$. In this paper, we propose a new way to test whether α_k is subject to the uniform distribution of $(0, 1)$.

Firstly, the theorem proved in Literature [4] is introduced: Suppose the population X is subject to uniform distribution in the interval (a, b) , and $X_{(1)}, X_{(2)}, \dots, X_{(n)}$ is the sequence statistics of X .

$$\text{The statistic } K_{ijk} = \frac{X_{(j)} - X_{(i)}}{X_{(k)} - X_{(i)}} \quad (1 \leq i < j < k \leq n)$$

obeys the Beta distribution $B(j-i, k-j)$ with the parameters of $j-i$ and $k-j$. Then, the distribution density of K_{ijk} can be expressed as:

$$f_{K_{ijk}}(z) = \frac{\Gamma(k-i)}{\Gamma(j-i)\Gamma(k-j)} z^{j-i-1} (1-z)^{k-j-1}$$

$$= \frac{(k-i-1)!}{(j-i-1)!(k-j-1)!} z^{j-i-1} (1-z)^{k-j-1},$$

$$0 < z < 1$$

According to the above theorem, the $\alpha_1, \alpha_2, \dots, \alpha_N$ is sorted from small to large to form a sequential sub-sample, denoted as $\alpha_{(1)} \leq \alpha_{(2)} \leq \dots \leq \alpha_{(N)}$, to construct the variable-scale differential quotient

$$\text{sequence statistic } Z_{ijk} = \frac{\alpha_{(j)} - \alpha_{(i)}}{\alpha_{(k)} - \alpha_{(i)}} \quad (1 \leq i < j < k \leq N).$$

So: if α_k obeys uniform distribution of $(0, 1)$, then the statistic Z_{ijk} obeys Beta distribution $B(j-i, k-j)$, and the distribution function of Z_{ijk} can be expressed as:

$$F_{Z_{ijk}}(z) = \int_0^z \frac{(k-i-1)!}{(j-i-1)!(k-j-1)!} t^{j-i-1} (1-t)^{k-j-1} dt$$

$$, \quad 0 < z < 1 \quad (2)$$

At this point, the test of whether α_k obeys the uniform distribution of $(0, 1)$ is transformed into the test of whether or not the statistic Z_{ijk} obeys the Beta distribution $B(j-i, k-j)$. Take the two-sided test as an example, given the significance level α , find the two-sided α -quantiles z_1 and z_2 corresponding to the distribution function of Z_{ijk} by using $F_{Z_{ijk}}(z_1) = 0.5\alpha$ and $F_{Z_{ijk}}(z_2) = 1 - 0.5\alpha$. If: $z_1 < Z_{ijk} < z_2$, the statistic Z_{ijk} obeys Beta distribution $B(j-i, k-j)$, so that α_k obeys uniform distribution of $(0,1)$, that is to

say both sub-samples x_1, x_2, \dots, x_n and y_1, y_2, \dots, y_m follow the same distribution; on the contrary, x_1, x_2, \dots, x_n and y_1, y_2, \dots, y_m two group of data not consistent, we can see that this group of simulation data is not credible.

4 Simulation date credibility assessment example

Assuming some index data obtained in the live test are: 8.7515, 8.8324, 8.5529, the corresponding index data obtained in the simulation test are: 8.7487, 8.6074, 8.4520, 8.9405, 8.7099, 8.8396, 8.8253, 8.7927, 8.6356, 8.6443, 8.7320, 8.6077, 8.6276, 8.9901, 8.7979, according to the above credibility assessment ideas, simulation data credibility assessment steps under small sample circumstance are:

(1) The simulation data were randomly divided into 5 groups, one of which was arbitrarily extracted as 8.7487, 8.8253 and 8.6276, and the two groups of sub-samples were sorted from small to large to form sequential sub-sample $x_{(k)}$.

(2) Considering the sufficiency of simulation information and the small sample feature of the live operational test data, the distribution of the fitting simulation sub-samples can be used to represent the population distribution $F(x)$. Therefore, by using the Lilliefors test [5] and the normfit fitting function, the population can be assumed to obey $N(8.7318, 0.1439^2)$ Normal distribution. Using the normcdf function to get the population distribution $F(x_{(k)})$ corresponding to the sequence sub-sample $x_{(k)}$, and then the corresponding α_k which was obtained by programming operation according to Equation (1) was ordered to form a sequential sub-sample $\alpha_{(k)}$, as shown in Table 1.

(3) The value of the variable-scale differential quotient sequence statistic Z_{ijk} is obtained from

$$Z_{ijk} = \frac{\alpha_{(j)} - \alpha_{(i)}}{\alpha_{(k)} - \alpha_{(i)}},$$

z_1, z_2 corresponding to each Z_{ijk} are obtained by using the betainv function according to Equation (2). If $z_1 < Z_{ijk} < z_2$, then the group of simulation sub-samples is credible; otherwise not credible. By means of MATLAB programming, the above calculating process is easy to achieve. Table 2 lists part of the Z_{ijk}, z_1, z_2 calculation results when the significance level $\alpha = 0.03$, from which we can see that the group of simulation sub-samples is credible.

(4) According to the above calculating steps, verify the credibility of the remaining 4 sets of simulation data.

Table 1. $x_{(k)}, \alpha_k, \alpha_{(k)}$ calculation results.

x_1	x_2	x_3	y_1	y_2	y_3
8.7515	8.8324	8.5529	8.7487	8.8253	8.6276
$x_{(1)}$	$x_{(2)}$	$x_{(3)}$	$x_{(4)}$	$x_{(5)}$	$x_{(6)}$
8.5529	8.6276	8.7487	8.7515	8.8253	8.8324
α_1	α_2	α_3	α_4	α_5	α_6
0.4925	0.4289	0.7393	0.4505	0.5152	0.1893
$\alpha_{(1)}$	$\alpha_{(2)}$	$\alpha_{(3)}$	$\alpha_{(4)}$	$\alpha_{(5)}$	$\alpha_{(6)}$
0.1893	0.4289	0.4505	0.4925	0.5152	0.7393

Table 2. Part of Z_{ijk}, z_1, z_2 calculation results.

	Z_{ijk}	z_1	z_2
1	0.9173	0.0150	0.9850
2	0.7903	0.0075	0.8775
3	0.8615	0.1225	0.9925
		
18	0.2780	0.2466	0.9950
19	0.2241	0.1225	0.9925
20	0.0920	0.0150	0.9850

5 Conclusion

This paper studies how to evaluate the credibility of the operational test simulation data under small sample circumstance when there is a small sample size of the live operational test data. By analyzing the characteristics and deficiencies of the existing research ideas and methods, a new judgment idea is proposed, which is not limited by the sample size. Using MATLAB programming can quickly determine the credibility of the simulation data, mainly including the following Step: ① Randomly group the simulation data; ② Solve the sequence sub-sample $\alpha_{(k)}$; ③ Construct the variable-scale differential quotient sequence statistic Z_{ijk} ; ④ Compare the value of Z_{ijk} and the values of z_1 and z_2 , and determine the credibility of the current group of simulation data; ⑤ Follow the above steps to test the rest of the groups.

References

1. Wang Kai, Zhao Dinghai, Yan Yaodong et al. Weapons and equipment operational test [M]. Beijing: National Defense Industry Press, 2012: 114-119

2. Tang Xuemei. Consistency Test Methods in Small Situation [J]. Systems Engineering and Electronics, 2001.23 (10): 66-80
3. Wang Hui, Pan Gao Tian, Zang Xingzhen et al. Theories and Methods of Compatibility Test on Normal Distribution with Few Samples [J] .Mathematics Practice and Theory, 2005.35 (3): 131-137
4. PAN Gao-Tian, PAN Feng, WANG Hui, et al. Researches On Test Theory and Method on Few Samples Test with Order Statistics [J] .Journal of Engineering Mathematics, 2005.22 (2): 301-306
5. Zhang Defeng. Probability and Mathematical Statistics using MATLAB [M]. Beijing: Mechanical Industry Press, 2010: 188-191