Defence Science Journal, Vol. 60, No. 3, May 2010, pp. 271-275

Success Probability Assessment Based on Information Entropy

Xuan Chen*, Hanyan Huang, and Zhengming Wang

National University of Defense Technology, Changsha, CO 410073, China
**E-mail: cxhueann@163.com*

ABSTRACT

The Bayesian method is superior to the classical statistical method on condition of small sample test. However, its evaluation results are not so good if subjective prior information is intervened. The success probability assessment about the success or failure tests of weapon products focussed in this paper, and a fusing evaluation method based on information entropy is proposed. Firstly, data from equivalent surrogate tests is converted into the prior information of an equivalent source by the information entropy theory. Secondly, the prior distribution of the success probability is identified via the Bootstrap method, and the posterior distribution is provided by the Bayesian method with the information of prototype tests in succession. Lastly, an example is given, and the results show that the proposed method is effective and valuable.

Keywords: Information entropy, equivalent surrogate test, equivalent source, Bayesian method, Bootstrap method

1. INTRODUCTION

The new-style weapons are usually costly, and a lot of geodesic equipments are needed with high precisions. The experiments about those weapons are usually enormous, complicated and with high investments. As a result, it is impossible to accomplish a great deal of prototype tests. So it is difficult to evaluate the performance of the weapons effectively just by data from a few prototype tests. Therefore, a lot of equivalent surrogate tests are usually put up to extend the information, including physical surrogate tests and simulations. Those equivalent surrogate tests are heterogeneous. It is obviously very important to seek advanced evaluation methods with all useful information adapting for weapons.

Both the classical statistical method and the Bayesian method can be used for estimating the success probability. However, on condition of small sample test, the classical statistical method is unfit, and the Bayesian method is preferred[1-5].

For the Bayesian method, it is a key problem to present the prior distribution by information from different kinds of tests. Different prior distributions of the success probability denoted as P, can be got from different types of heterogeneous information sources, and they should be fused into one integrated prior distribution logically. Then a series of conclusions can be derivable using the Bayesian theory. Generally speaking, the key of fusion is to identify the weights of different prior distributions logically[1,2].

Many relative studies have been done. A lot of researchers mainly put up linear models in terms of the distribution parameter to choose the weights. A few researchers put up non-linear models based on fuzzy operator theory[1-5]. However, if the built models were not so exact, the corresponding calculated weights would not be accurate. Consequently the evaluating results at last would not be good.

In success or failure tests, the prior distribution of the success probability P follows a β distribution in a homogeneous environment. Subjective experiential information, which takes distribution function $\pi(P)$ as $\beta(0,0)$, $\beta(1/2,1/2)$ or $\beta(1,1)$ via Jeffreys principles, is usually intervened to identify the distribution parameters[1,2,6,7]. However, when quantity of the samples are very small, the evaluating results are eventually various as $\pi(P)$ takes the three different values mentioned above. It is consequently very difficult to choose the correct distribution function $\pi(P)$ in applications.

This paper focusses on the evaluation of the success ratio and the information from equivalent surrogate tests is converted into the prior information in the equivalent tests based on the information entropy theory[8,9]. Then the prior distribution of success probability is identified by the Bootstrap method. At last the posterior distribution is provided by the Bayesian method with data from prototype tests.

2. HETEROGENEOUS INFORMATION CONVERSION

Due to the information entropy theory, the information is a decreasing, monotonous, and additive mapping of probability. The average information in terms of discrete variables is defined as following:

$$H = -\sum_{i=1}^{K} P_i \ln P_i \tag{1}$$

where, P_i is the happening probability in the i^{th} source,

Received 16 April 2009, Revised 14 September 2009

and K is the total number of sources.

Assume the number of independent heterogeneous sources is N. The respective success times is s_i and the respective failure times is f_i, then total number of the i^{th} kind of tests is $n_i = s_i + f_i$. The success probability of the i^{th} source is p_i, while the failure probability is $1-p_i$. As a result, the average information of every test provided in the i^{th} source is given by:

$$H_i = -[p_i \ln p_i + (1-p_i)\ln(1-p_i)] \tag{2}$$

Then, the information of all tests provided in the i^{th} source is given by $n_i H_i$, and the information of all tests provided in all sources is given by:

$$I = \sum_{i=1}^{N} n_i H_i = -\sum_{i=1}^{N} n_i [p_i \ln p_i + (1-p_i)\ln(1-p_i)] \tag{3}$$

All heterogeneous sources are to be converted into the prior information in an equivalent source. Assume that the number of the relative success times is s_0 and the number of the relative failure times is f_0, then total number of the equivalent tests is $n_0 = s_0 + f_0$. Let the success probability and the failure probability of the equivalent source be P and $1-P$, respectively. So the information provided in the equivalent source is given by

$$I' = -n_0[P\ln P + (1-P)\ln(1-P)] \tag{4}$$

Take the maximum likelihood estimator (MLE) $\widehat{p_i}$ of p_i, and $\widehat{P}$ of P as:

$$p_i = \widehat{p_i} = s_i / n_i, \quad P = \widehat{P} = s_0 / n_0 \tag{5}$$

Due to the principle that the total information should be equal, viz. $I' = I$. One obtains:

$$n_0 = \frac{\sum_{i=1}^{N}(s_i \ln s_i + f_i \ln f_i - n_i \ln n_i)}{P\ln P + (1-P)\ln(1-P)}$$
$$f_0 = n_0(1-P), \qquad s_0 = n_0 P \tag{6}$$

3. ESTIMATOR OF SUCCESS PROBABILITY IN EQUIVALENT SOURCE

There are four unknown variables, denoted as n_0, f_0, s_0, and P in Eqn (6). However, there are only three equations. The estimator of the success probability P in the equivalent source is therefore expected. Firstly, P is estimated just with physical surrogate sources, then it is computed with simulation sources only, at last both of them were incorporated to estimate P.

3.1 Physical Surrogate Sources

Physical surrogate tests include subsystems tests and components test of the weapons. They are usually put up to extend sources. For example, the reliability of weapons is evaluated. The estimated value of probability P would be different if the structural models of reliability were various. Denote the success probability computed in the physical surrogate sources as P_{phy}.

If the structural model of reliability were series-wound, P_{phy} should be expressed as following:

$$P_{\text{phy}} = \prod_{i=1}^{N} \widehat{p_i} = \prod_{i=1}^{N} \frac{s_i}{n_i} \tag{7}$$

If the structural model of reliability were shunt-wound, P_{phy} should be expressed as following:

$$P_{\text{phy}} = 1 - \prod_{i=1}^{N}(1-\widehat{p_i}) = 1 - \prod_{i=1}^{N} \frac{f_i}{n_i} \tag{8}$$

3.2 Simulation Sources

Simulation is under the principle that the model of it is similar to the real physical system. However, the model will not be strictly the same.

The simulated result would be much reliable if it is based on the standard of verification, validation and accreditation (VVA)[10]. However, if samples were very limited, the work for VVA would not be accomplished effectively. As a result, the simulated credibility should be defined for logical application of simulation sources.

The evaluating model about the simulated credibility is:

$$C_{sim} = \sum_i \eta_i B_i \tag{9}$$

where, C_{sim} is the simulated credibility, η_i is the weightiness coefficient of the i^{th} simulated part, which satisfies $\sum_i \eta_i = 1$, and B_i is the veracity coefficient of the i^{th} simulated part. η_i is usually decided by experts and professionals subjectively. B_i can be computed according to the reliability formula by prior information[2,11]. The simulated result is denoted as $U = (U_1, \ldots, U_{n_1})$ and the prototype result as $V = (V_1, \ldots, V_{n_2})$. The null hypothesis is denoted as H_0; here, U and V are in the same population. The alternative hypothesis is denoted by H_1; here, U and V are not in the same population. Denote A as accepting H_0, $\overline{A}$ as rejecting H_0, and the success probability computed with the simulated sources is denoted as P_{sim}. The credibility formula by prior information can be expressed as follows:

$$P_{sim} = P(H_0 \mid A) = \frac{1}{1 + \frac{(1-P(H_0))}{P(H_0)} \cdot \frac{\beta}{1-\alpha}} \tag{10}$$

where, α is the probability of the errors type I, β is the probability of the errors type II, and $P(H_0)$ is the happening probability of H_0. $P(H_0)$ is usually given by experts and professionals, according to fidelity of the i^{th} simulated part. The probabilities of the errors type I and the probabilities of the errors type II, denoted as α and β, are usually chosen by the aims and desires, and $\alpha = \beta$ is usually given in engineering application.

3.3 Total Sources of Tests

Total sources consist of physical surrogate sources and simulated sources. As described above, the success probability computed with physical surrogate sources is denoted as P_{phy}, and the success probability computed with simulated sources is denoted as P_{sim}. Then the success probability computed with total sources should be expressed as

$$\hat{P} = C_{\text{sim}} \cdot P_{\text{sim}} + (1 - C_{\text{sim}}) \cdot P_{\text{phy}} \tag{11}$$

4. BAYESIAN ESTIMATOR OF SUCCESS PROBABILITY

4.1 Prior Distribution of Success Probability

A Bernoulli variable is denoted as X, and the value of the variable is 0 or 1. Let $P\{X=1\}=p$, and let $P\{X=0\}=q=1-p$. Let $X_1,\cdots,X_{n_0}$ denote the historical sample, which are independent and with identical distribution (i.i.d). Then,

$$\hat{p} = \frac{1}{n_0}\sum_{i=1}^{n_0} X_i \tag{12}$$

$$T_{n_0} = \hat{p} - p \tag{13}$$

Take $\hat{q} = 1 - \hat{p}$ as distributed parameter of the new Bernoulli population, and resample the new population. That is to produce a series of random values of Bernoulli variable by computer, $X_1^*,\cdots,X_{n_0}^*$, and call them a regenerative sub-sample.

The average of regenerative sub-sample is:

$$\hat{p}^*_{Boot} = \frac{1}{n_0}\sum_{i=1}^{n_0} X_i^* \tag{14}$$

Then $R_n^* = \hat{p}^*_{Boot} - \hat{p}$ is denoted as the Bootstrap statistic. Their value will be different as the regenerative sub-samples are different. So let $X_1^{(j)},\ldots,X_{n_0}^{(j)}$ denote the j^{th} regenerative sub-sample, for $j=1,\ldots,N$, then one can get

$$R_n^{(j)*} = \hat{p}^{(j)*}_{Boot} - \hat{p},\ j = 1,\ldots,N \tag{15}$$

where N>>1. Consequently the experiential distributed function of R_n^* can be obtained. Take the distribution as the approach of distributed function of T_{n_0}. It can be proved that experiential distribution of R_n^* approaches to the distribution of T_{n_0} consentaneously.

After getting the distribution of T_{n_0}, and then $p = \hat{p} - T_{n_0}$, which is the distribution of p could be identified. Let it be the prior distribution of p and denote it as $\pi(p)$.

In success or failure tests, the prior distribution of p follows a β distribution, which is defined as $\pi(p) = \text{Be}(a,b)$. According to the Bootstrap method, the average $\bar{p}$ and the variance S^2 of the prior distribution can be obtained by

$$\bar{p} = \frac{1}{N}\sum_{j=1}^{N} p_j,\quad S^2 = \frac{1}{N-1}\sum_{j=1}^{N}\left(p_j - \bar{p}\right)^2 \tag{16}$$

where, p_j is the j^{th} Bootstrap estimate of the parameter p, $j=1,\ldots,N$.

Let the average $\bar{p}$ and the variance S^2 be equal to the expectation and the variance of β distribution. Hence

$$\begin{cases} \dfrac{a}{a+b} = \bar{p} \\ \dfrac{ab}{(a+b)^2(a+b+1)} = S^2 \end{cases} \tag{17}$$

Then the estimator of the parameter a and b can be solved by Eqn (17).

$$\begin{cases} \hat{a} = \bar{p}\left(\dfrac{(1-\bar{p})\bar{p}}{S^2} - 1\right) \\ \hat{b} = \left(1-\bar{p}\right)\left(\dfrac{(1-\bar{p})\bar{p}}{S^2} - 1\right) \end{cases} \tag{18}$$

4.2 Posterior Estimator of Success Probability

Suppose there are n prototype samples $x_1,\ldots,x_n$, where $x_i \sim f(x \mid p)$, $x_i (i=1,\ldots,n)$ is the i.i.d sample. Based on the Bayesian theory, the posterior distribution $\pi(p \mid X)$ of the success probability p can be obtained from the prior distribution $\pi(p)$ and the prototype samples $X = (x_1,\ldots,x_n)$ by the following formula:

$$\pi(p \mid X) = \frac{\pi(p) f(x \mid p)}{\int_{\Theta} \pi(p) f(x \mid p) dp} \tag{19}$$

where, Θ is the domain of p, and it is usually given as [0,1]. The point estimator and confidence interval of p are therefore easy to obtain.

Suppose that the number of total prototype tests times is m and the number of relative success times is n. Since the prior distribution of p is $\pi(p) = \text{Be}(a,b)$, it is concluded by formula (19) that the posterior distribution is $\pi(p \mid X) = \text{Be}(a+n, b+(m-n))$. Hence, the point estimator is:

$$\hat{p} = \frac{a+n}{a+b+m} \tag{20}$$

Given the confidence level δ, the confidence lower limit R_L of p is defined as

$$\int_{R_L}^{1} \pi(p \mid X) dp = \delta \tag{21}$$

Hence, R_L can be solved by the following formula:

$$\sum_{i=0}^{b+(m-n)} C_{a+b+m}^{i} R_L^{a+b+m-i} (1-R_L)^i = 1 - \delta \tag{22}$$

5. SIMULATION RESULTS AND ANALYSIS

The target capture probability is an essential criterion of radar detecting capacity. It will be taken as a evaluation example to show the effectiveness of the above mentioned method.

For some missiles, aiming at target on sea, it is impossible to evaluate the target capture probability effectively by data from very small number of prototype tests. Results of surrogate tests, including simulations, tests aiming at target on ground and tests with missile hung on airplane, are fused to evaluate target capture probability. Tests aiming at target on ground and tests with missile hung on airplane are all physical surrogate tests.

Factors affecting the target capture probability mainly consist of the target, the clutter environment, and the missile velocity. The compare tests aiming at target on ground to the prototype tests, the target, and the missile velocity are the same, but the clutter environment is different. The compare tests with missile hung on airplane to the prototype tests, the target and the clutter environment are the same but the missile velocity is different; The compare simulations to prototype tests, fidelity and similarity of the simulated model are considered.

Tests aiming at target on ground are regarded as subsystem tests, while tests with missile hung on airplane are component tests. Both of them are independent of each other, and the target capture probability can be calculated with sources of physical surrogate tests by Eqn (7).

While considering simulations, the sea clutter, the electromagnetic scatter of target, and the ballistic trajectory are needed to be simulated. Hence simulated credibility can be calculated with sources of simulations by Eqn (9). Simulated total launching times and relative success times in capturing target of physical surrogate tests, simulations and prototype tests are shown in Table 1.

Weightiness coefficients and veracity coefficients of the three simulated parts are shown in Table 2. The simulated credibility is $C_{sim} = 0.905$, with data simulated in Table 2. Taking $\alpha = \beta = 0.5$, hence the success probability computed in total sources is $\widehat{P} = 91.81$ % with data in Table 1 by Eqn (11). Total launching times and relative success times in the equivalent tests are afterwards $n_0 = 1006$, $s_0 = 924$ by Eqn (6).

Table 1. Target capturing results of all tests

	Total times	Success times
Tests aiming at target on ground	3	3
Tests with missile hung on airplane	20	18
Simulations	1000	920
Prototype tests	2	2

Table 2. Weightiness and veracity data of simulations

	Weightiness coefficient	Veracity coefficient (%)
Sea clutter	0.3	90
Electromagnetic scatter	0.3	85
Ballistic trajectory	0.4	95

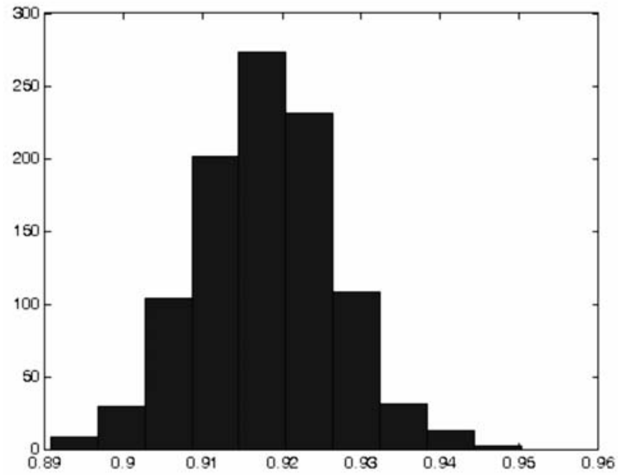

Figure 1. Prior distribution of target capture probability.

After computing prior information in an equivalent test, the prior distribution of p is shown in Fig. 1 by the Bootstrap method. Parameter a and b of prior distribution is a=896, b=80 by Eqn (18), with data in Fig. 1.

The results of the method proposed in this paper were compared with the results of the classical statistical method and the classical Bayesian method, the confidence level is set as $\delta = 0.95$:

(1) Consider the classical statistical method, that is to say, only the data of prototype tests can be used. The point estimator and the confidence lower limit of target capturing probability are $\widehat{p} = 100$ %, $R_L = 22.36$ %. It can be seen that the results are too conservative, since times of the prototype samples are too small and data of surrogate tests are not used.

(2) Consider the classical Bayesian method. The posterior distribution of target capturing probability is $\pi(p) = \text{Be}(943,82)$, with data in Table 1. The point estimator and the confidence lower limit of target capturing probability are $\widehat{p} = 92$ %, $R_L = 90.56$ %. It can be seen that the results are hazardous, since the heterogeneity between the surrogate tests and the prototype tests is not considered, and the data of prototype tests are inundated with the data of surrogate tests sequentially.

(3) Consider the method based on information entropy theory in this paper. After calculating the prior distribution $\pi(p) = \text{Be}(896,80)$, the posterior distribution of the target capturing probability is $\pi(p \mid X) = \text{Be}(898,80)$ via the Bayesian formula. The point estimator and the confidence lower limit of the target capturing probability are $\widehat{p} = 91.82$ %, $R_L = 90.33$ %. Since the prior distribution in the equivalent source is identified by the Bootstrap method, the subjective experiential factors are then not brought in. So the evaluated results are effective and logical.

When using the fusing evaluation method proposed in this paper, one can notice:

(1) In success or failure tests, the situation that the success probability is 1 or 0 sometimes appears. In this situation, the value ln 0 doesn't exist when Eqn (6) is calculated. In the actual calculation one needs to add a proper decimal fraction. It can be seen that the information is zero in the situation of none invalidation or none success.

(2) Estimator of the success probability p in the equivalent source involves two parts: physical surrogate tests and simulations. Success probability in physical surrogate tests, P_{phy}, has various estimators in different situations. Different situations are required to be considered and compute P_{phy}. In terms of simulations, weightiness coefficients and veracity coefficients of different simulated parts need experts and professionals to evaluate. Only when P_{phy} and C_{sim} were computed exactly one should ensure veracity of success probability p in the equivalent source.

6. CONCLUSIONS

In this paper, the data from multi-source heterogeneous tests is converted into the prior information of an equivalent source by the information entropy theory. Then the prior distribution of success probability is identified via the Bootstrap method, and the posterior distribution is given by the Bayesian formula with information of prototype tests in succession. The example shows that the proposed method is effective and valuable.

The method proposed is applicable in other field of success or failure tests.

ACKNOWLEDGEMENTS

This work was supported in part by the Natural Science Foundation of China (60974124) and the Spaceflight Footstone Technology Foundation Project in China.

REFERENCES

1. Zhang, Jinhuai & Liu, Qi. Bayesian test analysis methods. National University of Defense Technology Publishing Company, Changsha, 2007. pp. 3.
2. Tang, Xuemei; Zhang, Jinhuai; Shao, Fengchang & Li, Rong. Test analysis and evaluation of weapon systems in small-sample circumstances. National Defense Industry Publishing Company, Beijing, 2001. pp. 12.
3. Feng, Jing; Liu, Qi; Zhou, Jinglun & Dong, Chao. Correlation information fusion method and application in reliability analysis. *Syst. Eng. Electro.*, 2003, **25**(6), 682-84.
4. Wang, Jianxia & Guo, Bo. Bayes assesment for reliability of success or failure product based on fusing tests information of development phases. *Adv. Manuf. Manage.*, 2006, **25**(12), 31-32,41.
5. Liu, Han & Guo, Bo. Fusion of information of multiple sources in bayesian assessment for reliability of success or failure product. *J. Projectiles, Rockets, Missiles Guidance*, 2007, **27**(2), 338-41.
6. Giribone, R. & Valette, B. Principles of failure probability assessment (PoF). *Int. J. Pressure Vessels Piping*, 2004, **81**, 797-06.
7. Velarde, Luis Guillermo C.; Migon, Helio S. & Alcoforado, David A. Hierarchical Bayesian models applied to air surveillance radars. *Euro. J. Oper. Res.*, 2006, **184**(3), 1152-162.
8. Lia, Hui; Baoa, Yuequan & Ou, Jinping. Structural damage identification based on integration of information fusion and shannon entropy. *Mech. Syst. Signal Proc.*, 2008, **22**, 1427-440.
9. Sun, Youchao. Reliability synthesis assessment based on the information theory for sophisticated products. *J. Appli. Sci.*, 2001, **19**(2), 113-16.
10. Zhang, Wei & Wang, Xingren. Simulation credibility. *J. Syst. Simu.*, 2001, **13**(3), 312-14.
11. Huang, Hanyan; Duan, Xiaojun & Wang, Zhengming. A novel posterior-weighted bayesian estimation method considering the credibility of the prior information. *Acta Aeronautica Et Astronautic Sinica*, 2008, **29**(5), 1245-251.
12. Huang, Wei; Feng, Yunwei & Lv, Zhenzhou. Study on small scale sample test estimation method based on Bootstrap method. *Mech. Sci. Technol.*, 2006, **25**(1), 31-35.

Contributors

Mr Xuan Chen obtained his BS and MS both in Mathematics in 2003 and 2006, respectively, from College of Science, National University of Defense Technology, Changsha, China. His research areas include: test data processing, test design, and data fusion.

Ms Hanyan Huang obtained her BS from Wuhan University, China, in 2003 and MS from College of System Engineering, National University of Defense Technology, Changsha, China, both in System Engineering. Her research interests include: test data processing, experimental design, and data fusion.

Dr Zhengming Wang obtained his BS and MS, both in Mathematics in 1982 and 1986, respectively, from the Huazhong University of Science and Technology, Wuhan, China, and PhD in Control Science and Engineering in 1998 from National University of Defense Technology, Changsha, China. Since 2004, he is the Dean of the College of Science at National University of Defense Technology. He is the author or co-author of more than 90 papers and four books. His research interests include: image processing, test data processing, experimental design, and data fusion.