Original article
DOI: 10.21683/1729-2646-2019-19-3-12-20

Plan of tests with addition

Viktor S. Mikhailov, *D.I. Mendeleev Central Research and Design Institute of Chemistry and Mechanics, Russian Federation, Moscow*

Viktor S. Mikhailov

Abstract. *It is common practice to estimate the values of dependability indicators (point estimation). Normally, the probability of no-failure (PNF) is used as the dependability indicator. Due to economic reasons, determinative dependability tests of highly dependable and costly products involve minimal numbers of products, expecting failure-free testing (acceptance number Q = 0) or testing with one failure (Q = 1), thus minimizing the number of tested products. The latter case is most interesting. By selecting specific values of the acceptance number and number of tested products, the tester performs a preliminary estimation of the planned PNF, while selecting Q = 1 the tester minimizes the risks caused by an unlikely random failure. However, as the value Q grows, the number of tested products does so as well, which makes the testing costly. That is why the reduction of the number of products tested for dependability is of paramount importance.* **Preparation of the plan of tests with addition.** *We will consider binomial tests (original sample) with addition of one product (oversampling) to testing in case of failure of any of the initially submitted products. Testing ends when all submitted products have been tested with any outcome (original sampling and oversampling). Hereinafter it is understood that the testing time is identical for all products. Testing with the acceptance number of failures greater than zero (Q > 0) conducted with addition allows reducing the number of tested products through successful testing of the original sample. The* **Aim** *of the paper consists in preparing and examining PNF estimates for the plan of tests with addition.* **Methods of research of dependability indicator estimates.** *Efficient estimation is based on the integral approach formulated in [6, 8-10]. The integrative approach is based on the formulation of the rule of efficient estimate selection specified on the vertical sum of absolute (or relative) biases of estimates selected out of a certain set based on the distribution law parameter, where, in our case, n is the number of products initially submitted to testing.* **Criterion of selection of efficient estimation for PNF.** *The criterion of selection of an efficient estimate of the probability of failure (or PNF) at a set of estimates is based on the total square of absolute (or relative) bias of the mathematical expectation of estimates* $E\hat{\theta}(n,k,m)$ *from probability of failure p for all possible values of p, n.* **Conclusions.** *PNF estimates for the plan of tests with addition was prepared and examined. For the case n > 3, the PNF estimate* $\hat{P}(n,k,m) = 1-\hat{p}(n,k,m) = 1-(k+m)/(n+k)$ *in comparison with the implicit estimate* $\hat{V}(n,k,m) = 1-\hat{v}(n,k,m)$ *is bias efficient. Testing with the acceptance number of failures greater than zero (Q > 0) conducted with addition allows reducing the number of tested products through successful testing of the original sample. Estimates* $\hat{p}_2$, $\hat{w}_2$ *and* $\hat{w}_3$ *are unbiassed and, as a consequence, bias efficient for the cases n = 2 and n = 3 respectively.*

Keywords: *Bernoulli scheme, test plan, point estimation, probability of no-failure, efficient estimate, mean time to failure*

For citation: *Mikhailov VS. Plan of tests with addition. Dependability 2019; 3: 12-20 p. DOI: 10.21683/1729-2646-2019-19-3-12-20*

Introduction

It is common practice to estimate the values of dependability indicators (point estimation). Normally, the probability of no-failure (PNF) is used as the dependability indicator. Due to economic considerations, determinative dependability tests of highly dependable and costly products involve minimal numbers of products, expecting failure-free testing (acceptance number Q=0) or testing with one failure (Q=1), thus minimizing the number of tested products. The latter case is most interesting. By selecting specific values of the acceptance number Q and number of tested products, the tester performs a preliminary estimation of the planned PNF, while selecting Q=1 the tester minimizes the risks caused by an unlikely random failure. However, as the value Q grows, the number of tested products does so as well, which makes the testing costly. That is why the reduction of the number of products tested for dependability is of paramount importance.

Preparation of the plan of tests with addition

We will consider binomial tests (original sample) [1, 2] with addition of one product (oversampling) to testing in case of failure of any of the initially submitted products. Testing ends when all submitted products have been tested with any outcome (original sampling and oversampling). Hereinafter it is understood that the testing time is identical for all products.

Testing with the acceptance number of failure greater than zero (Q>0) conducted with addition allows reducing the number of tested products through successful testing of the original sample.

The Aim of the paper

The aim of the paper consists in preparing and examining PNF estimates for the plan of tests with addition.

Preparation and examination of PNF estimates for the plan of tests with addition

Let n be the number of tested products of the same type initially submitted to testing, and let $R = r$ be the number of failed products that includes k failures from n products initially submitted to testing and m failures from k products repeatedly submitted to testing, i.e. $r=k+m$. Then, the number of tested products will be $N=n+k$. Let failures be independent events, then the probability of r failures during tests (hereinafter, $P_n(R=r)$) is easily expressed with a generating function. Let us apply properties of the generating function [3].

The generating function (hereinafter, $\psi_R(z)$) is a mathematical expectation of an exponential function of type z^R, i.e. for the test plan with addition [3]: $\psi_R(z) = Ez^R = \sum_{i=0}^{2n} z^i P_n(R=i)$.

For the case when the original sample consists of one product, the generating function will be [3]:

$$\psi_R(z) = Ez^R = \sum_{i=0}^{2n} z^i P_n(R=i) = q + qpz + p^2z^2.$$

Then, for the case when original sample consists of n products, the generating function will be [3]:

$$\psi_{n;R}(z) = \left(q + qpz + p^2z^2\right)^n.$$

The probability of zero failures during testing of the original sample with volume n [3]:

$$P_n(R=0) = \psi_{n;R}(z) = \left(q + qpz + p^2z^2\right)^n |_{z=0} = q^n.$$

The mathematical expectation of the random value R can be calculated through the expression [3]: $ER = \psi^{(1)}_{n;R}(z=1)$ is the first derivative.

And the probability of getting exactly r failures can be calculated through the expression [3]:

$$P_n(R=r) = \psi^{(r)}_{n;R}(z=0)/r!.$$

Let us construct the derivative of the generating function:

$$\psi^{(1)}_{n;R}(z) = n\left(q + qpz + p^2z^2\right)^{n-1}\left(2p^2z + pq\right),$$

out of which follows that the average number of failures during tests will be

$$ER = \psi^{(1)}_{n;R}(z=1) = n\left(q + qp + p^2\right)^{n-1}\left(2p^2 + pq\right) = np(1+p).$$

Then, the probability of one failure during tests can be calculated by the formula:

$$P_n(R=1) = \psi^{(1)}_{n;R}(z=0) = nq^{n-1}pq = npq^n$$

The construction of derivatives of the higher orders is very complicated, and therefore it is not demonstrated in this paper.

The obtained results are not the best option for calculations, therefore, let us construct a more convenient formula for the probability of exactly r failures during tests that is obtained from the following construction procedure ($n \geq k \geq m; r = k + m \leq 2n$):

$$P_k(m) := C_k^m p^m q^{k-m};$$

$$P_n(k) := C_n^k p^k q^{n-k} \sum_{m=0}^{k} P_k(m) = C_n^k p^k q^{n-k},$$

where q=1–p, p is the probability of failure, C_n^k is the number of k combinations of n elements.

$$P_n(k,m) := P_n(k)P_k(m) = C_n^k C_k^m p^{k+m} q^{n-m};$$

$$P_n(r=0) = P_n(k=0, m=0) = q^n;$$

$$P_n(r=1) = P_n(k=1, m=0);$$

$$P_n(r=2; r\le n) = P_n(k=1, m=1) + P_n(k=2, m=0);$$

$$P_n(r=3; r\le n) = P_n(k=2, m=1) + P_n(k=3, m=0);$$

$$P_n(r=4; r\le n) = P_n(k=2, m=2) + \\ + P_n(k=3, m=1) + P_n(k=4, m=0);$$

$$P_n(r=5; r\le n) = P_n(k=3, m=2) + \\ + P_n(k=4, m=1) + P_n(k=5, m=0);$$

$$P_n(r=6; r\le n) = P_n(k=3, m=3) + P_n(k=4, m=2) + \\ + P_n(k=5, m=1) + P_n(k=6, m=0);$$

$$P_n(r=7; r\le n) = P_n(k=4, m=3) + P_n(k=5, m=2) + \\ + P_n(k=6, m=1) + P_n(k=7, m=0);$$

$$\ldots$$

$$P_n(R=r) = \sum_{k=0}^{n} \sum_{m:m+k=r, m\le k} P_n(k,m);$$

$$\ldots$$

$$P_n(r=2n) = P_n(k=n, m=n) = p^{2n}.$$

From the construction logic we obtain the required formula for the probability of exactly r failures:

$$P_n(R=r) = \sum_{k=0}^{n} \sum_{m:m+k=r, m\le k} P_n(k,m),$$

where$r = k + m = 0, 1, 2, \ldots, 2n$; $k = 0, 1, 2, \ldots, n$; $m : m + k = r, m \le k$.

Through the calculation of probability $P_n(k=x,m=y)=P_n(k=x)$ $P_n(m=y)$, where $x, y = 0, 1, 2, \ldots, n$ and $P_n(R=r)$ it is easy to obtain the probability function of the plan of tests with addition:

$$P_{n\sum}(k\le x, m\le y) = \sum_{k=0}^{x} \sum_{m:m+k\le x+y, m\le k, m\le y} P_n(k,m), \quad (1)$$

which on the entire set of events $r=k+m=0,1,2,\ldots, 2n$ should be equal to one. Let us verify this fact.

The probability function on the entire set of events can be represented as the sum of the products of each component of the primary polynomial by polynomial, where polynomials have binominal coefficients:

$$P_{n\sum}(n,n) = \sum_{r=0}^{2n} P_n(r) = \sum_{k+m=0}^{2n} P_n(k) P_k(m) =$$

$$= \sum_{k+m=0}^{2n} C_n^k p^k q^{n-k} C_k^m p^m q^{k-m} = q^n + C_n^1 p^1 q^{n-1} \sum_{m=0}^{1} C_1^m p^m q^{1-m} + \ldots +$$

$$+ C_n^k p^k q^{n-k} \sum_{m=0}^{k} C_k^m p^m q^{k-m} + \ldots + p^n \sum_{m=0}^{n} C_n^m p^m q^{n-m} =$$

$$= \sum_{k=0}^{n} C_n^k p^k q^{n-k} = 1,$$

or:

$$P_{n\sum}(n,n) = \sum_{k=0}^{n} P_n(k) = \sum_{k=0}^{n} C_n^k p^k q^{n-k} = 1.$$

An expression for ER can also be found in a simpler way. The average number of tested products during tests with addition consists of the number of products that were originally submitted to testing and the average number of failed products that were originally submitted to testing, i.e. $N=n+np$. Then, the average number of failed products during tests with addition will be:

$$E(R,n) = Np = E(k,n) + E(m,n) = \\ = np + np * p = (n+np)p = np(1+p).$$

Let us note that the probability $P_n(k,m)$ defines the test results (k,m), therefore, as an estimate of parameter p it is recommended to choose an estimate that defines the maximum probability $P_n(k,m)$.

Let us solve the classical problem of identification of function maximum

$$b(r,p,k,n) = P_n(k,m) = C_n^k C_k^m p^{k+m} q^{n-m}$$

with respect to p. For that, let us take the logarithm for the function $b(r,p,k,n)$, let us take the derivative with respect to the variable p, set the result to zero and solve an equation with respect to variable p. The resulting estimate $\hat{p} = r/(n+k) = r/(n+r-m)$ determines the maximum of function $b(r,p,k,n)$. Let us consider the properties of the resulting estimate $\hat{p} = r/(n+k)$ and the PNF estimate, as a consequence

$$\hat{P} = 1 - \hat{p} = 1 - r/(n+k) = (n-m)/(n+k).$$

Let $k+m=r>1$, then for various $k_1>k_2$, $m_1<m_2$ the following inequality will be true

$$\hat{p}(k_1,m_1) = \frac{r}{n+k_1} < \hat{p}(k_2,m_2) = \frac{r}{n+k_2}, \quad (2)$$

i.e. the dependability of the controlled batch of products (PNF: $\hat{P}(k_1,m_1) = 1 - \hat{p}(k_1, m_1)$) according to the test of a sample, in which the number of products failed during test k_1 was greater than in the sample of a comparable batch k_2 with the same number r of failures will always be higher $\hat{P}(k_1,m_1) > \hat{P}(k_2,m_2)$ than in a comparable batch of products. In other words, when comparing the results of two finally formed samples (with equal numbers of failures), the priority in dependability is given to the products, whose failures were primarily within the original sample, and not oversampling. In this case, oversampling enables remediation after unsuccessful initial tests. This is the advantage of the test plan with addition.

Unbiassed estimate calculation

Let us determinate the mathematical expectation of the estimate $\hat{p}(n;k,m) = r/(n+k)$:

$$E\left(\hat{p}\left(n;k,m\right)\right)=\sum_{r=0}^{2n}\frac{r}{n+k}P_{n}\left(r\right).$$

It can be proved that estimate $E(\hat{p}(n;k,m))$ in general is biased. To prove that, a particular case will suffice.

Let us determine the mathematical expectation of estimate $\hat{p}(n=1) = r/(1+k)$:

$$n=1:E(\hat{p}(n=1))=\sum_{r=0}^{2}\frac{r}{1+k}P_1(r)=\ 0*P_1(k=0,m=0)+$$
$$+\frac{1}{2}P_1(k=1,m=0)+1*P_1(k=1,m=1)=$$
$$=\frac{1}{2}pq+p^2=0,5(p+p^2).$$

Therefore, estimate $\hat{p}(n=1) = r/(1+k)$ is biased. Estimate $\hat{p}(n=1)$ can be presented in the following form:

$$\hat{p}\left(n=1\right)=\frac{r}{1+k}\equiv\begin{cases}0\ ,r=0,k=0,m=0;\\ \frac{1}{2},r=1,k=1,m=0;\\ 1\ ,r=2,k=1,m=1.\end{cases}$$

By equating the mathematical expectation of unknown estimate $\hat{w}_1(n=1;k,m)$ to parameter *p*, it is easy to obtain an unbiased estimate of probability of failure $\hat{w}_1$ for the case $n=1; p_0, p_1, p_2$ are unknown probabilities:

$$E\left(\hat{w}_1\right)=\sum_{r=0}^{2}\frac{r}{1+k}\hat{w}_1P_1\left(r\right)=p_0\left(1-p\right)+p_1\left(p-p^2\right)+p_2p^2=p;$$

$$p^0:p_0\,p^0=p_0*1=0=>p_0=0;p^1:p_1p^1=p=>p_1=1;$$

$$p^2:-p^2p_1+p^2p_2=0=>p_2=p_1=1;$$

$$\hat{w}_1\equiv\begin{cases}0\quad ,r=0,k=0,m=0;\\ 1\quad ,r=1,k=1,m=0;\\ 1\quad ,r=2,k=1,m=1.\end{cases}$$

An unbiased estimate is an indicator function, i.e. in case of failures estimate $\hat{w}_1$ is equal to one, if otherwise, this estimate is equal to zero. The option when *n*=1 is practically not interesting, because it coincides with the binominal plan, therefore, it will not be considered in this paper.

Let us determine the mathematical expectation for $\hat{p}(n=2) = r/(2+k)$

$$n=2:E(\hat{p}(n=2))=\sum_{r=0}^{4}\frac{r}{2+k}P_2(r)=\ 0*P_2\ (k=0,m=0)+$$
$$+(1/3)P_2(k=1,m=0)+(2/3)P_2(k=1,m=1)+$$
$$+(1/2)P_2(k=2,m=0)+(3/4)P_2(k=2,m=1)+1*$$
$$*P_2(k=2,m=2)=0*q^2+(1/3)2p^1q^2+(2/3)2p^2q+$$
$$+(1/2)p^2q^2+(3/4)2q^3q+1*p^4=2p\ (1-p)(1/3-(1/3)p+$$
$$+(2/3)p+(1/4)p-(1/4)p^2)+(3/2)p^3-(3/2)p^4+p^4=$$
$$=2p(1-p)(1/3+(7/12)p-(1/4)p^2)+(3/2)p^3-$$
$$-(3/2)p^4+p^4=\left(\frac{2}{3}\right)p+\left(\frac{7}{6}\right)p^2-\left(\frac{1}{2}\right)p^3-\left(\frac{2}{3}\right)p^2-\left(\frac{7}{6}\right)p^3+$$
$$+\left(\frac{1}{2}\right)p^4+\left(\frac{3}{2}\right)p^3-\left(\frac{3}{2}\right)p^4+p^4=\left(\frac{2}{3}\right)p+\left(\frac{1}{2}\right)p^2-\left(\frac{1}{6}\right)p^3\ ;$$

$$p=0,5:E(\hat{p}(n=2))=1/3+1/8-1/(6*8)=21/48\ .$$

Therefore, estimate $\hat{p}(n=2) = r/(2+k)$ is biased. Estimate $\hat{p}$(n=2) can be presented as:

$$\hat{p}(n=2)\equiv\begin{cases}0,\quad r=0,k=0,m=0;\\ 1/3,r=1,k=1,m=0;\\ 2/3,r=2,k=1,m=1;\\ 1/2,r=2,k=2,m=0;\\ 3/4,r=3,k=2,m=1;\\ 1,\quad r=4,k=2,m=2.\end{cases}$$

Let us note that for the results $\hat{p}(r=2,k=1,m=1)=2/3$ and $\hat{p}(r=2,k=2,m=0)=1/2$ the dependability of the controlled batch of products, in which some products in the sample failed during initial test, is higher than in the products whose failures occurred during repeated test and with the same number of failures. That corresponds to the property of estimate $\hat{p} = r/(n+k)$, expressed by formula (2).

It is easy to obtain an unbiased estimate $\hat{s}_2$ for parameter *p*:

$$\hat{s}_2\equiv\begin{cases}0,\quad r=0,k=0,m=0;\\ 1/2,r=1,k=1,m=0;\\ 5/8,r=2,k=1,m=1;\\ 6/8,r=2,k=2,m=0;\\ 7/8,r=3,k=2,m=1;\\ 1,\quad r=4,k=2,m=2.\end{cases}$$

For this purpose the mathematical expectation of the supposed unbiased estimate with unknown probabilities p_{ik} must be equated to parameter *p* and necessary transformations must be carried out:

$$E(\hat{p}(n=2)=\sum_{r=0}^{4}\hat{p}(n=2)P_2(r)=[p_{00}=0;p_{22}=1]=p_{00}q^2+$$
$$+p_{10}2pq^2+p_{11}2p^2q+p_{20}p^2q^2+p_{21}2p^3q+p^4=2p_{10}p-$$
$$-2p_{10}2p^2+2p_{10}p^3+2p_{11}p^2-2p_{11}p^3+p_{20}p^2-2p_{20}p^3+$$
$$+p_{20}p^4+p_{21}2p^3-p_{21}2p^4+p_{22}p^4=2p_{10}p-4p_{10}p^2+$$
$$+2p_{11}p^2+p_{20}p^2+2p_{10}p^3-2p_{11}p^3-2p_{20}p^3+p_{21}2p^3+$$
$$+p_{20}p^4-p_{21}2p^4+p_{22}p^4=p.$$

For this equation to be true, the coefficients with different degrees of parameter p must be equal to zero, with the exception of the first degree where the coefficient must be equal to one:

$$p^1:2p_{10}p^1=p=>p_{10}=1/2;$$
$$p^2:-4p_{10}p^2+2p_{11}p^2+p_{20}p^2=0=>2p_{11}+p_{20}=2=>$$
$$=>2p_{11}=2-p_{20};$$
$$p^3:2p_{10}p^3-2p_{11}p^3-2p_{20}p^3+2p_{21}p^3=0=>2p_{11}+2p_{20}-2p_{21}=$$
$$=1=>2-p_{20}+2p_{20}-p_{20}-1=1=>p_{20}=6/8=>p_{11}=5/8;$$
$$p^4:p_{20}p^4-p_{21}2p^4+p_{22}p\,^\wedge 4=0=>[p_{22}=1]:2p_{21}-p_{20}=$$
$$=p_{22}=>2p_{21}=p_{20}+1=>p_{21}=7/8.$$

This heterogeneous system of linear equations is always resolvable and has an infinite set of similar solutions (the number of variables is greater than the number of equations):

$$p_{00}=0;p_{10}=1/2;p_{11}=5/8;p_{20}=6/8;p_{21}=7/8;p_{22}=1.$$

Let us note that the failure probabilities must satisfy the slack inequality $0\le p_{ij}\le 1$. Let us also point out that, in practice, for two controlled batches of products with the same number of failures in the generated samples for the results $p_{20}=6/8$ and $p_{11}=5/8$ the dependability of the first controlled batch of products $1-p_{20}=1-6/8=2/8$, for which some products in the original sample and oversampling failed only during initial tests k=2, m=0, is lower than for products of the second controlled batch $1-p_{11}=1-5/8=3/8$, where failures occurred during repeated tests in oversampling as well. This result contradicts the property (see formula (2)) of the biased estimate $\hat{p}(r=k+m,k,m)=r/(n+k)$ and makes it difficult to choose an efficient estimate.

Further, in order to avoid contradictions when looking for new estimates of the failure probability, it is necessary to take into account that the values of estimates for the same number of failures do not depend on the fact, in which sample (original or additional) the failures occurred. Therefore, this principle of looking for new estimates of the failure probability $\hat{w}(n;k,m)$ can be presented as follows:

$$\hat{w}(k\_1+m\_1=r,k\_1,m\_1)=$$
$$=\hat{w}(k\_2+m\_2=r,k\_2,m\_2), \qquad (3)$$

i.e. we reject the property estimate $\hat{p}$ expressed by formula (2).

Similarly to the above reasoning, let us demonstrate the method of finding new estimates:

$$\hat{w}_2(0):=p_0;\ \hat{w}_2(1):=p_1;\ \hat{w}_2(2):=p_2;\ \hat{w}_2(3):=p_3;\ \hat{w}_2(4):=p_4$$

$$E(\hat{w}_2)=\sum_{r=0}^{4}\hat{w}_2(r)P_2(r)=p=p_0q^2+p_12pq^2+p_22p^2q+$$
$$+p_2p^2q^2+p_32p^3q+p_4p^4=2p_1p-2p_12p^2+2p_1p^3+$$
$$+2p_2p^2-2p_2p^3+p_2p^2-2p_2p^3+p_2p^4+p_32p^3-p_32p^4+$$
$$+p_4p^4\,2p_1p-4p_1p^2+2p_2p^2+p_2p^2+2p_1p^3-2p_2p^3-$$
$$-2p_2p^3+p_32p^3+p_2p^4-p_32p^4+p_4p^4.$$

In order for this equality to be true, the coefficients with different degrees should be equal to zero, with the exception of the first degree, where the coefficient should be equal to one:

$$p^0:p_0p^0=p_0*1=0=>p_0=0;$$

$$p^1:2p_1=1=>p_1=1/2;$$

$$p^2:-4p_1p^2+2p_2p^2+p_2p^2=0=>-2+3p_2=0=>p_2=2/3;$$

$$p^3:2p_1p^3-2p_2p^3-2p_2p^3+p_32p^3=$$
$$=0=>1-8/3+2p_3=0=>p_3=5/6;$$

$$p^4:p_2p^4-p_32p^4+p_4p^4=$$
$$=0=>2/3-10/6+p_4=0=>p_4=1;$$

$$p_0=0;p_1=1/2;p_2=2/3;p_3=5/6;p_4=1.$$

This heterogeneous system of linear equations is always solvable and has only one solution (the number of variables 2*n is equal to the rank (number of linearly independent equations) [5]), which will be estimate $\hat{w}_2$!

Similarly to the previous example (case n=2), let us determine the mathematical expectation of estimate $\hat{p}(n=3)=r/(3+k)$:

$$n=3:\ E(\hat{p}(n=3))=\sum_{r=0}^{6}\frac{r}{3+k}P_3(r).$$

After all the required manipulations (they are not presented due to complicated expressions) the following result will be obtained: estimate $\hat{p}(n=3)=\sum_{r=0}^{6}\frac{r}{3+k}$ is biased. Estimate $\hat{p}(n=3)=\sum_{r=0}^{6}\frac{r}{3+k}$ is presented as follows:

$$\hat{p}(n=3)=\sum_{r=0}^{6}\frac{r}{3+k}\equiv\begin{cases}0, r=0, k=0, m=0;\\ 1/4, r=1, k=1, m=0;\\ 1/2, r=2, k=1, m=1;\\ 2/5, r=2, k=2, m=0;\\ 3/5, r=3, k=2, m=1;\\ 4/5, r=4, k=2, m=2;\\ 1/2, r=3, k=3, m=0;\\ 2/3, r=4, k=3, m=1;\\ 5/6, r=5, k=3, m=2;\\ 1, r=6, k=3, m=3.\end{cases}$$

Let us determine an unbiased estimate of failure probability for the case n=3 ($\hat{w}_3(r)$), using the principle expressed by the formula (3). The probability values of this estimate are determined through its mathematical expectation that should be equal to the estimated parameter p:

$$\hat{w}_3(0):=p_0; \hat{w}_3(1):=p_1; \hat{w}_3(0):=p_2; \hat{w}_3(3):=p_3;$$
$$\hat{w}_3(4):=p_4; \hat{w}_3(5):=p_5; \hat{w}_3(6):=p_6;$$

$$E(\hat{w}_3(r))=p=p_0q^3+p_1 3pq^3+3p_2(p^2q^2+p^2q^3)+$$
$$+p_3(6p^3q^2+p^3q^3)+p_4(3p^4q+3p^4q^2)+p_5 3p^5q+$$
$$+p_6p^6=3p_1(p-3p^2+3p^3-p^4)+3p_2(p^2-2p^3+p^4)+$$
$$+3p_2(p^2-3p^3+3p^4-p^5)+6p_3(p^3-2p^4+p^5)+$$
$$+p_3(p^3-3p^4+3p^5-p^6)+3p_4(p^4-p^5)+$$
$$+3p_4(p^4-2p^5+p^6)+3p_5(p^5-p^6)+p_6p^6$$

$$p^0: p_0p^0=p_0*1=0; p_0=0;$$

$$p^1: 3p_1=1; p_1=1/3;$$

$$p^2: -9p_1+3p_2+3p_2=0 => 6p_2=3 => p_2=1/2;$$

$$p^3: 9p_1-6p_2-9p_2+6p_3+p_3=0 => 6p_2+$$
$$9p_2-6p_3-p_3=3 => p_3=9/14;$$

$$p^4: -3p_1+3p_2+9p_2-12p_3-3p_3+3p_4+3p_4=$$
$$=0 => 12p_2-15p_3+6p_4=1 => p_4=65/84;$$

$$p^5: -3p_2+6p_3+3p_3-3p_4-6p_4+3p_5=0 => p_5=75/84;$$

$$p^6: -p_3+3p_4-3p_5+p_6=0 => p_6=1;$$

$$\hat{w}_3(r)\equiv\begin{cases}0, r=0;\\ 1/3, r=1;\\ 1/2, r=2;\\ 9/14, r=3;\\ 65/84, r=4;\\ 75/84, r=5;\\ 1, r=6.\end{cases}$$

A similar search for unbiased estimates for cases n=4 and n=5 was unsuccessful, because the obtained results of the probability values exceeded 1, which is not acceptable. Therefore, for n>3 the construction of an unbiased estimate according to the rule $\hat{p}(k_1+m_1=r, k_1, m_1)=\hat{p}(k_2+m_2=r, k_2, m_2)$ is problematic!

Let us introduce a new term: let the estimate of failure probability (hereinafter, $\hat{v}$) center the probability function $P_{n\sum}$ relative to the limits of its values. This means that intervals $[0;\hat{v}]$ and $[\hat{v};1]$ of values of such estimates with the probability 0.5 cover the estimated parameter p. Such estimates will be called centered. Let us note that for some test plans centered estimates are close to efficient estimates [6, 8]. In this case, the centered estimate v is calculated from the following expression (replacing p with v in the formula (1)):

$$P_{n\sum}(k\le x, m\le y, \hat{v})=\sum_{k=0}^{x}\sum_{m:m+k\le x+y, m\le k, m\le y}P_n(k, m, \hat{v})=0,5.$$

For the solution for this equation to exist be unique, it is necessary to verify the monotonicity of $P_{n\sum}$ with respect to variable p [1, 7]. It should be reminded that $P_n(k,m):=C_n^k C_k^m p^{k+m} q^{n-m}, r=k+m.$

Taking thee derivative of $P_{n\sum}$ to the parameter p, the results will be the following:

$$(P_{n\sum}(k\le x, m\le y, p))'_p=$$
$$=\sum_{k=0}^{x}\sum_{m:m+k\le x+y, m\le k, m\le y}C_n^k C_k^{r-k}(rp^{r-1}q^{n-r+k}-(n-r+k)p^r q^{n-r+k-1})$$

Due to the complexity of the obtained expression, it is not possible to prove or dispose of the monotonicity of $P_{n\sum}$. However, it is possible for the most interesting cases as r=0, r=1 и r=2. Let us consider these cases:

$$r=0:\ (P_{n\sum}(n, p, k=0, m=0))'_p=$$
$$=C_n^0 C_0^{0-0}\ [(0+0)p^{-1}q^n-np^0q^{n-1}]=-nq^{n-1}<0;$$

$$r=1:\ (P_{n\sum}(n, p, k=1, m=0))'_p=$$
$$=C_n^1C_1^0\left[(1+0)p^0q^n-npq^{n-1}\right]-C_n^0C_0^0nq^{n-1}=$$
$$=nq^n-n^2pq^{n-1}-nq^{n-1}=nq^{n-1}(1-p-np-1)=$$
$$=-pn(n+1)q^{n-1}<0;$$

$$r=2:\ (P_{n\sum}(n, p, k=1, m=1))'_p=C_n^1C_1^1[(1+1)pq^{n-1}-$$
$$-(n-1)p^2q^{n-2}]+C_n^1C_1^0[(1+0)p^0q^n--npq^{n-1}]-$$
$$-C_n^0C_0^0nq^{n-1}=2npq^{n-1}-n(n-1)p^2q^{n-2}+nq^n-n^2pq^{n-1}-$$
$$-nq^{n-1}=npq^{n-2}(2(1-p)-(n-1)p)-pn(n+1)q^{n-1}=$$
$$=npq^{n-2}(2-p-np)-pn(n+1)q^{n-1}=npq^{n-2}(2-n-1)\le 0$$

Table 1. The values of LCB of parameter *p* for different scopes of tests (in horizontal direction) and failure events (in vertical direction) if γ=0,8

	n	1	2	3	4	5	6	7	8
k=0	*m*=0	0.199	0.105	0.071	0.054	0.043	0.036	0.031	0.027
k=1	*m*=0	0.445	0.287	0.212	0.168	0.139	0.119	0.104	0.092
k=1	*m*=1	1	0.445	0.287	0.212	0.168	0.139	0.119	0.104

Table 2. The values of UCB of parameter *p* for different scopes of tests (in horizontal direction) and failure events (in vertical direction) if *a*=0,2

	n	1	2	3	4	5	6	7	8
k=0	*m*=0	0.800	0.552	0.415	0.331	0.275	0.235	0.205	0.182
k=1	*m*=0	0.894	0.710	0.582	0.488	0.422	0.370	0.330	0.297
k=1	*m*=1	1	0.894	0.710	0.582	0.488	0.422	0.370	0.330

Table 3. The values of the centered estimate $\hat{v}$ for different scopes of tests (in horizontal direction) and failure events (in vertical direction)

	n	1	2	3	4	5	6	7	8
k=0	*m*=0	0.292	0.206	0.159	0.129	0.108	0.094	0.082	0.074
k=1	*m*=0	0.707	0.5	0.384	0.313	0.264	0.226	0.201	0.179
k=1	*m*=1	1	0.707	0.5	0.384	0.313	0.264	0.226	0.201

$$r=2:\ (P_{n\sum}(n,p,k=2,m=0))'_p = C_n^2 C_2^0[(2+0)pq^n -$$
$$-np^2q^{n-1}\ ] + C_n^1 C_1^0[(1+0)p^0q^n - np^1q^{n-1}\ ] - C_n^0 C_0^0 nq^{n-1} =$$
$$= n(n-1)pq^n - 0,5n^2(n-1)p^2q^{n-1} + nq^n - n^2pqq^{n-1} -$$
$$-nq^{n-1} = n(n-1)pq^{n-1}(1-p-0,5np) - pn(n+1)q^{n-1} =$$
$$= npq^{n-2}(0,5np-n) \le 0.$$

Therefore, for cases when *r*=0, *r*=1, *r*=2 probability function $P_{n\sum}$ monotonically decreases with the increasing parameter *p* and, therefore, the centered estimate $\hat{v}$ of parameter *p* for the test plan with addition is unique.

The centered estimate defines the lower (upper) confidence boundary (hereinafter referred to as LCB (UCB)) of the interval of the unknown parameter *p* with the confidence probability γ=0,5 or significance level α=1–γ=0,5. On the other hand, any estimate of LCB (UCB) of the interval of unknown parameter *p* can be interpreted as a point estimate of parameter *p* with a strong bias (downward bias is for LCB and upward bias is for UCB). Unidirectional LCB (hereinafter referred to as $\hat{p}_L$) and UCB (hereinafter referred to as $\hat{p}_U$) of the interval with unknown parameter *p* with confidence probability γ=1–α are calculated in accordance with the following formulas:

$$P_{n\sum}(x,y,\hat{p}_L)=\gamma,\ P_{n\sum}(x,y,\hat{p}_U)=\alpha.$$

The boundaries of the central confidence interval are calculated in accordance with the following formulas [4]:

$$P_{n\sum}(x,y,\hat{p}_L)=1-\alpha/2,\ P_{n\sum}(x,y,\hat{p}_U)=\alpha/2\ .$$

Tables 1, 2 and 3 show the values of LCB, UCB of parameter *p* and values of the centered estimate $\hat{v}$ for the most realistic scopes of tests and failure events.

Let us formulate a criterion for choosing an efficient estimate of failure probability (or PNF) and construct – on the basis of the formulated criterion – an improved (and biased) estimate of failure probability (and, therefore, the estimate of the PNF) for the test plan with addition for *n*>3 and choose the efficient one among the proposed estimates.

Research methods for estimating dependability indicators

The search for efficient estimates is based on the integral approach described in [6, 8-10]. The integral approach is based on construction of the rule for choosing an efficient estimate $\hat{\theta}_0(n;k,m)$ specified on the sum of the absolute (or relative) bias of estimates of $\hat{\theta}_0(n;k,m)$, selected from a certain set, from the parameter of the distribution law, where in this case *n* is the number of products initially put up for testing.

Criterion for choosing and efficient estimate for PNF

The criterion for choosing an efficient estimate of the probability of failure (or PNF) over the set of estimates of $\hat{\theta}_0(n;k,m)$ is based on the total square of absolute (or relative) biases of mathematical expectations of estimates of $E\hat{\theta}(n;k,m)$ from the probability of *p* failure for all possible values*p*, *n*.

In order to select an efficient estimate of the probability of failure (or PNF) the concept of an absolutely efficient estimate by bias and parameter *p* variation within $0\le p\le 1$

are required. In order to obtain the final result, the functional (hereinafter referred to as $L(\hat{\theta}(n;k,m))$ on the limited set $n_1 \le n_i \le n_j$, $i = 1,\dots, j$ is constructed [6, 8–10] as a criterion for efficient estimate $\hat{\theta}(n;k,m)$:

$$L(\hat{\theta}(n;k,m)) = \frac{1}{j} \sum_{n_1 \le n_i \le n_j} \int_0^1 (E\hat{\theta}(n_i;k,m) - p)^2 \, dp \quad (4)$$

Estimate $\hat{\theta}_0(n;k,m)$ that minimizes the functional $L(\hat{\theta}(n;k,m))$ on a given set of estimates is an efficient bias estimate on a given set of biased estimates. Among estimates that similarly minimize functional $L(\hat{\theta}(n;k,m))$, an estimate that has the minimal mean-square deviation (classical definition of an efficient estimate [1]) is to be selected. We will call this estimate more efficient in comparison with the selected ones.

Selecting estimates with minimal deviation involves constructing a functional (hereinafter referred to as $D(\hat{\theta}(n;k,m)$) by summarizing mathematical expectations of squares of relative deviations of estimates of $\hat{\theta}(n;k,m)$ from parameter p for all possible values p, n [6, 8-10]:

$$D(\hat{\theta}(n;k,m)) = \frac{1}{j} \sum_{n_1 \le n_i \le n_j} \int_0^1 E(\hat{\theta}(n_i;k,m) - p)^2 \, dp \quad (5)$$

An estimate that provides zero to functional $L(\hat{\theta}(n;k,m)$)=0 (unbiased estimate) and minimizes functional $D($ $\hat{\theta}(n;k,m))$ will be called absolutely biased.

Let us limit the scope tests as $0 \le n \le 10$, which is the cost limit for highly reliable and complicated products. Then, formula (4) will be as follows:

$$L(\hat{\theta}(n;k,m)) = \frac{1}{10} \sum_{1 \le i \le 10} \int_0^1 (E\hat{\theta}(n_i;k,m) - p)^2 \, dp.$$

And formula (5) will be presented as:

$$D\left(\hat{\theta}(n;k,m)\right) = \frac{1}{10} \sum_{1 \le i \le 10} \int_0^1 E\left(\hat{\theta}(n_i;k,m) - p\right)^2 dp.$$

Table 4 shows the results of the substitution into functional $L(\hat{\theta}(n;k,m))$ and $D(\hat{\theta}(n;k,m))$ in accordance with formulas (4) and (5) of the following estimates of failure probability $\hat{\theta}$: $\hat{p}, \hat{p}_2, \hat{w}_2, \hat{w}_3, \hat{v}$. The calculations were carried out with the step of $dp = 10^{-3}$.

Table 4 shows that for options $n>3$ estimate $\hat{p}$ has a minimal bias compared to estimate $\hat{v}$. $\hat{p}_2$, $\hat{w}_2$ and $\hat{w}_3$ estimates are unbiased and, as a result, are efficient for options $n=2$ and $n=3$ respectively.

Table 4 shows that estimate $\hat{v}$ has a slight advantage over estimate $\hat{p}$ as regards minimal deviation of its values from parameter p. Therefore, the estimate $\hat{p} = (k+m)/(n+k)$ can be taken as a desired efficient bias estimate among the proposed ones.

Let us note that the variation of the step of summation changes the functional result, but does not change the result of estimates comparison.

Example. Products are part of a redundant piece of equipment. It is required to make a point estimate of PNF products according to the binominal reliability tests. While planning determinative dependability tests, the tester, when calculating sample volume (n=6), took into account only one failure (Q=1), minimizing the risk of the improbable unpredictable failure. In this case, the predicted PNF value was $\hat{P} = 1 - 1/n = 5/6 = 0{,}83$ that corresponds to the requirements of the technical specifications (PNF should be at least 0.83) for the product. Given that during tests the failure of product is unlikely, it was decided to carry out the dependability tests with addition to reduce the costs. During the test two outcomes are possible: no failure and one failure (as planned). In case of no failure there is no need for oversampling. Let us consider these options:

1) No-failure tests. No-failure tests with addition:

$$\hat{P} = 1 - \hat{p}(n=5, k=0, m=0) =$$
$$= 1 - r/(n+k) = 1 - 0/(5+0) = 1;$$

$$\hat{V} = 1 - \hat{v}(n=5, k=0, m=0) = 1 - 0{,}108 = 0{,}892.$$

One-sided LCB of PNF as $n=5, \gamma=1-\alpha=1-0{,}2=0{,}8$ was (see Table 2)

$$\hat{P}_н(n=5, r=0) = 1 - \hat{p}_в(n=5, k=0, m=0) =$$
$$= 1 - 0{,}275 = 0{,}725.$$

Binominal tests with one failure:

$$\hat{P}(n=6, r=0) = 1 - r/n = 1 - 0/6 = 1.$$

One-sided LCB of PNF as $n=6, r=0, \gamma=0{,}8$ (calculated according to the Clopper-Pearson equation [2]) was $\hat{P}_L(n=6, r=0) = (1-\gamma)^{\frac{1}{6}} = 0.764$.

2) Tests with one failure. Tests with addition and with one failure:

$$\hat{P} = 1 - \hat{p}(n=5, k=1, m=0) =$$
$$= 1 - r/(n+k) = 1 - 1/(5+1) = 0{,}83;$$

Table 4. Results of substitution of the proposed estimates of failure probability into functionals $L(\hat{\theta}(n;k,m))$ and $D(\hat{\theta}(n;k,m))$

Functional	$\hat{p}_2(n=2)$	$\hat{w}_2(n=2)$	$\hat{w}_3(n=3)$	$\hat{p}\ (n>3)$	$\hat{v}(n>3)$
$L(\hat{\theta}(n;k,m))$	$2.6 \cdot 10^{-33}$	$2.6 \cdot 10^{-33}$	$5.1 \cdot 10^{-33}$	$2 \cdot 10^{-4}$	$1.51 \cdot 10^{-3}$
$D(\hat{\theta}(n;k,m))$	0.0687	0.0418	0.0418	0.0187	0.0164

$$\hat{V} = 1 - \hat{v}(n = 5, k = 1, m = 0) = 1 - 0,264 = 0,736.$$

One-sided LCB of PNF as n=5, $\gamma=1-\alpha=1-0,2=0,8$ was (see Table 2)

$$\hat{P}_L(n = 5, r = 1) = 1 - \hat{p}_U(n = 5, k = 1, m = 0) = \\ = 1 - 0.422 = 0.588.$$

Binominal tests with one failure: $\hat{P} = 1 - r / n = 1 - 1/6 = 0,83.$

One-sided LCB of PNF as $n = 6, r = 1, \gamma = 0,8$ (calculated according to the Clopper-Pearson equation [2]) was $\hat{P}_L(n = 6, r = 1) = 0.578$.

Conclusions

PNF estimates for the plan of tests with addition was prepared and examined. For the case of n>3, the PNF estimate $\hat{P}(n,k,m) = 1 - \hat{p}(n,k,m) = 1 - (k+m)/(n+k)$ in comparison with the implicit estimate $\hat{V}(n,k,m) = 1 - \hat{v}(n,k,m)$ is bias efficient.

Testing with the acceptance number of failure greater than zero (Q>0) conducted with addition allows reducing the number of tested products through successful testing of the original sample.

Estimates $\hat{p}_2, \hat{w}_2, \hat{w}_3$ are unbiased and, as a consequence, bias efficient for the cases n=2 and n=3 respectively.

References

[1] Borovkov A.A. Matematicheskaya statistika [Mathematical statistics]. Novosibirsk: Nauka; 1997 [in Russian].

[2] Gnedenko B.V., Beliaev Yu.K., Soloviev A.D. Matematicheskie metody v teorii nadezhnosti [Mathematical methods in the dependability theory]. Moscow: Nauka; 1965 [in Russian].

[3] Krupkina T.V. Teoriya veroyatnostey i matematicheskaya statistika. Chast 2. Elektronnyy kurs lektsiy [Probability theory and mathematical statistics. Part 2. Electronic series of lectures]. Krasnoyarsk: Siberian Federal University; 2011 [in Russian].

[4] Shulenin V.P. Matematicheskaya statistika. Chast 1. Parametricheskaya statistika: ouchebnik [Mathematical statistics. Part 1. Parametric statistics: a textbook]. Tomsk: Izdatelstvo NTL; 2012 [in Russian].

[5] Kostrikin A.I. Vvedenie v algebru. Chast I. Osnovy algebry: uchebnik dlya vuzov [Introduction to algebra. Part 1. Basic algebra: textbook for higher educational institutions. Moscow: MCCME; 2004 [in Russian].

[6] Mikhailov V.S. Implicit estimates for the NBτ test plan. Reliability and quality of complex systems 2018;1(21):64-71 [in Russian].

[7] Fikhtengolts G.M. Kurs differentsialnogo i integralnogo ischisleniya. Tom 1 [Course of differential and integral calculus. Volume 1]. Moscow: Nauka; 1969 [in Russian].

[8] Mikhailov V.S., Yurkov N.K. Estimates of reliability indicators for fault-free tests conducted according to the binomial plan. Reliability and quality of complex systems 2018;4(24):29-39 [in Russian].

[9] Mikhailov V.S. Efficient estimation of mean time to failure. Dependability 2016;4:40-42.

[10] Mikhailov VS. Estimation of the gamma-percentile life for the binomial test plan. Dependability 2019; 2:18-21.

About the author

Viktor S. Mikhailov, Lead Engineer, D.I. Mendeleev Central Research and Design Institute of Chemistry and Mechanics, Russian Federation, Moscow, e-mail: Mvs1956@list.ru

Received on: 14.04.2019