COMPARISON OF SELECTED STRATEGIES OF STATISTICAL QUALITY CONTROL

Kazimierz Worwa*

* Faculty of Cybernetics, Military University of Technology, Gen. S. Kaliskiego 2. st., Warsaw, 00-908, Poland, Email: kworwa@wat.edu.pl

Abstract The paper presents the results of a comparative analysis of the two methods of statistical quality control of products. That two methods differ from each other mainly by a way of determining a subset of products subjected to inspection. The first of the two methods, hereinafter referred to as the completely random method, is to draw a certain number of products for control of the entire products population. The second method, called the method of drawing from partitions, is a pre-split across the specific products population into smaller parts, called partitions, and then create a representation (sample) of the products subject to control by drawing the products from each partition. The main result of this study is to determine the conditions for which the effectiveness of quality control methods with drawing from partitions is not less than the efficiency of the method with fully random drawing, whereby the efficiency criterion is the likelihood of the event, consisting in the fact that there is at least one incorrect product among the drawn products. The considerations were illustrated by numerical examples, designed to compare the effectiveness of the analyzed methods of quality control.

Paper type: Research Paper

Published online: 31 July 2015
Vol. 5, No. 4, pp. 337-346

ISSN 2083-4942 (Print)
ISSN 2083-4950 (Online)

Keywords: *Statistical Quality Control, Control Strategy, Control Effectiveness*

1. INTRODUCTION

The subject of consideration in this study is to evaluate the effectiveness of the two selected methods of quality control of products manufactured en masse, as a result of repetitive manufacturing (mass production). In practice, the number of products intended to control can be very large, and therefore control of all products would be very expensive and time consuming, and often impossible for technical reasons. In such cases the most frequently control method used in practice is sampling method, involving the controlling only a limited number of products, drawn from the entire products population. Methods for assessing the quality of the products, which are not subject to control their entire party and only certain random selected representation (sample), are called statistical methods of quality control (Hamrol & Mantura, 2002).

The idea of statistical methods for quality control of products is based on inference about the quality of the whole products population on the basis of the analysis of a some sample that is created randomly (PN-EN ISO 9000:2001, 2001). This sample is a specific representation of the entire products population.

The purpose of sampling is to obtain information about the quality of the whole population by examining so small the number of its representatives as possible. Random inspection of the quality of products is used in practice when testing the whole product population is too expensive or time-consuming, or if the quality control is a destructive. As a result of sampling it can be determined whether the percentage of defective products is acceptable. Acceptable level of defective products in a specified party is generally defined in terms of delivery and receipt and agreed between the manufacturer and the customer.

The opposite of statistical control is a control, which includes all manufactured products. Such control is time-consuming and is in practice used only for the products manufactured individually or in small series.

The two methods of statistical products quality control considered in the paper differ from each other mainly by a way of determining a subset of products subjected to inspection. The first of the two methods, hereinafter referred to as the completely random method, is to draw a certain number of products for control from the entire products population. The second method, called the method of drawing from partitions, is a pre-split across the products population into smaller parts, called partitions, and then create a representation (sample) of the products subject to control by drawing the products of each partition. The method of subdividing entire batch production of the product on the partitions may depend on the specifics of the production process, including for example its organization. A single partition can be a group of products, e.g. from one production line, or produced by a certain set of production (e.g. a single employee) or as a result of work of one shift of production, etc.

2. DESCRIPTION OF THE QUALITY CONTROL METHODS

To assess the effectiveness of the statistical products quality control methods that are compared we can use the results of some work towards comparison of selected methods for determining a set of test cases in the process of software testing, using the random methods (Weyuker & Jeng, 1991), (Chen & Yu, 1996), (Worwa, 2009).

Let ***M*** denote the set of products produced in a some production period, and its size is M, $M>0$. Products that are a result of conducted quality control have been recognized as incorrect product will be called incorrect product, and their number will be denoted by K, $0 \le K \le M$. Other products, i.e. products that successfully would pass inspection will be referred to the correct products, and their number will be denoted by L. Of course, there is $K+L=M$.

Let p is the probability of the event consisting in the fact that as a result of drawing single product from the set ***M*** we get an incorrect product. According to earlier designations that probability is defined as follows

$$p = K / M \ . \tag{1}$$

For further consideration the probability p will be called the production error factor.

For purposes of comparison, the effectiveness of two statistical products control methods that are considered we assume that the total number of products selected for inspection is N, $N>0$.

In the case of the method of drawing with partitions the set of all products ***M*** is divided into I subsets called partitions, denoted by $\boldsymbol{M}_i$, $i=1, 2, ..., I$. It is assumed that a subset of $\boldsymbol{M}_i$ has size M_i and contains either K_i incorrect product and L_i correct products, wherein occurs $K_i + L_i = M_i$, $i = 1,2,...,I$. The error production rate and the *i*-th batch is defined as

$$p_i = K_i / M_i \, , \ \ i = 1,2,...,I \ . \tag{2}$$

Number of products drawn from the i-th batch is N_i, $N_i \ge 1$.To ensure the proper conditions for the comparison of the effectiveness of both statistical products control methods that are analyzed it is assumed that there is

$$\sum_{i=1}^{I} N_i = N \ . \tag{3}$$

Equation (3) means that the comparison of the effectiveness of both statistical products control methods will be carried out for the same total number of products selected for inspection. Assuming that the subsets $\boldsymbol{M}_i$ are disjoint we can write

$$\sum_{i=1}^{I} M_i = M \ , \quad \sum_{i=1}^{I} L_i = L \ , \quad \sum_{i=1}^{I} K_i = K \ . \tag{4}$$

In further considerations it is assumed that drawing products into the control is performed without replacement in accordance with a uniform distribution, i.e. random selection of each product that is to be controlled is equally likely. This means in particular that if the product is randomly chosen from the set of ***M*** or ***Mi***, the probability that the incorrect product will be drawn is p or p_i, respectively.

As the main criterion for comparing the effectiveness of two statistical products control methods that are considered will be considered the likelihood of the event, consisting in the fact that there is at least one incorrect product among the drawn products.

This probability will be denoted respectively by P_r for the fully random drawing, and P_p for the the drawing from partitions. These probabilities can be determined as follows (Weyuker & Jeng, 1991):

$$P_r(N) = 1-(1-p)^N \tag{5}$$

and

$$P_p(\boldsymbol{N}) = 1-\prod_{i=1}^{I}(1-p_i)^{n_i} , \tag{6}$$

wherein $\boldsymbol{N} = (N_1, N_2, ..., N_I)$, $I > 1$, is the vector of number of tests3 drawn from the ***i***-th subset of the manufactured products (partition).
It is worth noting that there is only one partition, i.e. for $I=1$ we have $P_p = P_r$. Similarly, if $p_1 = p_2 = ... = p_k = p$ we obtain

$$P_p(\boldsymbol{N}) = 1-\prod_{i=1}^{I}(1-p_i)^{N_i} =$$

$$= 1-\prod_{i=1}^{I}(1-p)^{N_i} = 1-(1-p)^{N_1+N_2+\dots+N_I} = 1-(1-p)^N = P_r(N),$$

which means that the probability P_r and P_p are equal. It means that the effectiveness of both analyzed methods of statistical quality control is the same.

It would seem that in terms of the number of detected incorrect products the more expensive method of drawing products from partitions should be more efficient than the cheaper method of completely random drawing. In reality, however, differences in the effectiveness of both methods are very small, and very often cheaper method of completely random drawing can be even more effective than the more expensive method of drawing from partitions.

It is easy to note that the method of drawing products from partitions will be better than fully random method, if the majority of randomly selected products will come from the partitions having a larger error production rate. This requirement does not guarantee, however, the occurrence of that effect, what is illustrated in Example 1 (Worwa, 2009). In this example, there are three drawing cases are considered, each of which is based on three partitions ($I=3$), wherein in each case the total number of randomly selected products and the total number of incorrect products are the same.

For the first case, the method of drawing from partitions draw is better than fully random method, i.e. $P_r < P_p$. in the second case is exactly the opposite, i.e. $P_r > P_p$, while in the third case, both probabilities are equal, i.e. $P_r = P_p$.

It is clear that in practice it is impossible to have a priori accurate information on the distribution of incorrect products in particular partitions that could be used in a method for drawing from partitions.

Despite the absence of such information using the results obtained in the work (Weyuker & Jeng, 1991) it is possible to determine how to divide the set of all product onto subsets of products (partitions)(partitions), which ensures that the method of drawing from partitions will not be worse than completely random method. The authors of the cited work showed that if $M_1 = M_2 = \cdots = M_I$ and $N_1 = N_2 = \cdots = N_I$, to $P_p \geq P_r$, then the method of drawing from partitions will not be worse than completely random method. Moreover, if incorrect products are evenly distributed among the partitions, i.e. $K_1 = K_2 = \cdots = K_I$ then $P_p = P_r$. Thus, the method of drawing from partitions never be worse than completely random method if the partition will have the same cardinality, i.e. $M_1 = M_2 = \cdots = M_I$ and from each of them will be randomly selected the same number of products, i.e. $N_1 = N_2 = \cdots = N_I$. This effect is illustrated in Example 2 (Worwa, 2009). This condition is a sufficient condition to guarantee that the method of drawing from partitions would be more effective than the completely random method. Unfortunately, the fulfillment of this condition in practice can be very difficult.

Example 1

I=3											
	M_i		K_i		N_i		p_i		p	P_r	P_p
1	M_1	1200	K_1	10	N_1	10	p_1	0,0083	0,01	0,1821	0,1912
	M_2	500	K_2	7	N_2	7	p_2	0,014			
	M_3	300	K_3	3	N_3	3	p_3	0,01			
Total		2000		20		20					
2	M_1	1200	K_1	11	N_1	13	p_1	0,0092	0,01	0,1821	0,1725
	M_2	500	K_2	7	N_2	3	p_2	0,014			
	M_3	300	K_3	2	N_3	4	p_3	0,0067			
Total		2000		20		20					
3	M_1	1200	K_1	12	N_1	13	p_1	0,01	0,01	0,1821	0,1821
	M_2	500	K_2	5	N_2	3	p_2	0,01			
	M_3	300	K_3	3	N_3	4	p_3	0,01			
Total		2000		20		20					

Example 2

	I=3										
	M_i		K_i		N_i		p_i		p	P_r	P_p
1	M_1	700	K_1	10	N_1	50	p_1	0,0143	0,0100	0,7785	0,7788
	M_2	700	K_2	7	N_2	50	p_2	0,0100			
	M_3	700	K_3	4	N_3	50	p_3	0,0057			
Total		2100		21		150					
2	M_1	700	K_1	7	N_1	50	p_1	0,0100	0,0100	0,7785	0,7785
	M_2	700	K_2	7	N_2	50	p_2	0,0100			
	M_3	700	K_3	7	N_3	50	p_3	0,0100			
Total		2100		21		150					

In the Example 2 two cases are considered, each of which contains three partitions. In the first case $M_1 = M_2 = M_3$ and $N_1 = N_2 = N_3$. Then there is $P_p \geq P_r$. In the latter case an additional condition is satisfied $K_1 = K_2 = K_3$, and then the probability values P_r and P_p are identical. In the paper (Chen & Yu, 1996) made the following generalization of this condition indicated at earlier work (Weyuker & Jeng, 1991): if the partitions are disjoint than method of drawing from partitions will be more efficient than the fully random method, if the number of products that are drawn from each partition is proportional to the cardinality of these partitions. Formally, this means that if $I \geq 2$ and $N_1/M_1 = N_2/M_2 = \cdots = N_I/M_I$ than $P_p \geq P_r$. This effect is illustrated in Example 3 (Worwa, 2009), where *I=3* and $N_1/M_1 = N_2/M_2 = \cdots = N_I/M_I = 0{,}06$. For each of the three cases in Example 3 the total number of tests is the same *N=2100* and the total number of incorrect products $K = K_1 + K_2 + K_3$ is different for each case. It will be noticed that in each of the three cases considered in Example 3, the method of drawing from partitions is more effective than fully random method. We can also see that the increase of the number of incorrect test $K = K_1 + K_2 + K_3$ entails an increase in the number of incorrect products triggers an increase the values of both probabilities P_r and P_p.

It is worth noting that the values of probabilities P_p in Example 3 are only slightly greater than the probabilities P_r. The reasons for this effect will be explained in chapter 3 of this paper.

Example 3

	I=3								p	P_r	P_p
	M_i		K_i		N_i		p_i				
1	M_1	500	K_1	10	N_1	30	p_1	0,0200	0,0095	0,6996	0,7019
	M_2	300	K_2	7	N_2	18	p_2	0,0233			
	M_3	1300	K_3	3	N_3	78	p_3	0,0023			
Total		2100		20		126					
2	M_1	500	K_1	10	N_1	30	p_1	0,0200	0,0143	0,8371	0,8377
	M_2	300	K_2	10	N_2	18	p_2	0,0333			
	M_3	1300	K_3	10	N_3	78	p_3	0,0077			
Total		2100		30		126					
3	M_1	500	K_1	20	N_1	30	p_1	0,0400	0,0286	0,9742	0,9747
	M_2	300	K_2	20	N_2	18	p_2	0,0667			
	M_3	1300	K_3	20	N_3	78	p_3	0,0154			
Total		2100		60		126					

3. THE BEST AND WORST CASES FOR THE DRAWING FROM PARTITIONS METHOD

In the work (Worwa, 2009) there was defined the following conditions, the fulfillment of which guarantee to maximize the probability specified by (6):

1. If $I > 1$, $\sum_{i=1}^{I} N_i = N$ and the partitions are numbered in such a way that $p_1 \le p_2 \le ... \le p_I$, then there is

$$P_p(N) = 1 - \prod_{i=1}^{I}(1-p_i)^{N_i} \le 1 - \prod_{i=1}^{I-1}(1-p_i)(1-p_I)^{N-I+1}, \tag{7}$$

i.e.

$$P_p(1,1,...1,N-I+1) = \max_{\boldsymbol{N}} P_p(\boldsymbol{N}). \tag{8}$$

2. If $I > 1$, $\sum_{i=1}^{I} N_i = N$ and the partitions are numbered in such a way that $p_1 \ge p_2 \ge ... \ge p_I$, then there is

$$P_p(\boldsymbol{N}) = 1 - \prod_{i=1}^{I}(1-p_i)^{N_i} \geq 1 - \prod_{i=1}^{I-1}(1-p_i)(1-p_I)^{N-I+1}, \qquad (9)$$

i.e.

$$P_p(1,1,...1,N-I+1) = \min_{\boldsymbol{N}} P_p(\boldsymbol{N}) \qquad (10)$$

where $\boldsymbol{N} = (N_1, N_2, ..., N_I)$.

The correctness of the above statements illustrate Examples 4 and 5 (Worwa, 2009).

Example 4

				N_1	N_2	N_3	P_p	P_r
				1	1	28	0,9900	
N	M	K	p	1	2	27	0,9887	
30	1200	60	0,05	1	3	26	0,9871	
				1	4	25	0,9854	
K_1	10			1	5	24	0,9833	
K_2	20					...		
K_3	30			10	12	8	0,8592	
				10	13	7	0,8398	
M_1	400			10	14	6	0,8178	
M_2	600			10	15	5	0,7928	0,7854
M_3	200			10	16	4	0,7644	
				10	17	3	0,7321	
p_1	0,025			10	18	2	0,6953	
p_2	0,033					...		
p_3	0,150			26	2	2	0,6504	
				26	3	1	0,6025	
				27	1	2	0,6474	
				27	2	1	0,5990	
				28	1	1	0,5956	

Example 4 illustrates the best and the worst case of the probability P_p for $I=3$, $N=30$ and $p_1 \leq p_2 \leq p_3$. According to equations (9) and (10) the probability P_p reaches a maximum for ***N****=(1, 1, 28)* and minimum for ***N****=(28, 1, 1)*, respectively.

It is noteworthy that, for ***N****=(10, 15, 5)* is $N_1/M_1 = N_2/M_2 = N_3/M_3 = 0{,}025$, which means compliance with the conditions laid at work (Chen &Yu, 1996). The row that meets this requirement has been shaded. In accordance with previous observations that condition ensures that there is $P_p \geq P_r$. However, it could be noticed that the probability value P_p is only slightly greater than the probability

P_r. Example 4 shows that there is $P_p(10,15,5)=0{,}7928$, while $P_r(30)=0{,}7854$. At the same time, however, we can see that the probability $P_p(10,15,5)=0{,}7928$ is significantly less than the probability $P_p(1,1,28)=0{,}9900$. In turn, the case of ***N*** = *(28, 1, 1)* is the best case of probability value P_p. In fact for this case is $P_p(1,1,28)=0{,}5956$.

Example 5

		N_1	N_2	P_p	P_r	N_1/M_1	N_2/M_2
		1	99	0,8836		0,0002	0,0066
		2	98	0,8830		0,0004	0,0065
		3	97	0,8824		0,0006	0,0065
		4	96	0,8817		0,0008	0,0064
I	2	5	95	0,8811		0,0010	0,0063
K_1	80		…			…	
K_2	320	21	79	0,8703		0,0042	0,0053
M_1	5000	22	78	0,8696		0,0044	0,0052
M_2	15000	23	77	0,8689		0,0046	0,0051
		24	76	0,8681		0,0048	0,0051
p_1	0,0160	25	75	0,8674	0,8674	0,0050	0,0050
p_2	0,0213	26	74	0,8667		0,0052	0,0049
		27	73	0,8660		0,0054	0,0049
		28	72	0,8652		0,0056	0,0048
		29	71	0,8645		0,0058	0,0047
		30	70	0,8638		0,0060	0,0047
			…			…	
		31	69	0,8630		0,0062	0,0046
		96	4	0,8050		0,0192	0,0003
		97	3	0,8039		0,0194	0,0002
		98	2	0,8029		0,0196	0,0001
		99	1	0,8018		0,0198	0,0001

Example 5 characterizes, respectively, the best and the worst case of the probability P_p for $I=2$, $p_1 \le p_2$ and $N = N_1 + N_2 = 100$. According to equations (9) and (10) the probability P_p reaches a maximum for ***N***=*(1, 99)*, and minimum for ***N***=*(99, 1)*, respectively. It is worth noting that for ***N***=*(25, 75)*occurs $N_1/M_1 = N_2/M_2 = 0{,}005$. The row that corresponds to this condition has been shaded.

3. CONCLUSION

The main result of this study is to determine the conditions for which the effectiveness of quality control methods with drawing from partitions is not less than the efficiency of the method with fully random drawing, whereby the efficiency criterion has been values of probabilities P_p and P_r, which are defined by (5) and (6), respectively. The considerations were illustrated by numerical examples (Examples 1-5), designed to compare the effectiveness of the analyzed methods of quality control.

REFERENCES

Chen T.Y. & Yu Y.T., (1996), "A more general sufficient condition for partition testing to be better than random testing", Information Processing Letters, Vol. 57, pp. 145-149.

Hamrol A. & Mantura W., (2002), Zarządzanie jakością. Teoria i praktyka. PWN, Warszawa.

PN-EN ISO 9000:2001, (2001), Systemy zarządzania jakością. Podstawy i terminologia. Polski Komitet Normalizacyjny, Warszawa.

Weyuker E. J. & Jeng B., (1991), "Analyzing partition testing strategies", IEEE Transactions on Software Engineering, Vol. SE-17, pp. 703-711.

Worwa K., (2009), "A comparison analysis of testing strategies based on subdomain testing and random testing", Proceedings of the Fourth International Conference on Dependability of Computer Systems DepCos – Relcomex, pp. 115-122.

BIOGRAPHICAL NOTES

Kazimierz Worwa is an Associate Professor at the Cybernetics Faculty of Military Technical University in Warsaw. He is deputy dean for education. His research interests include software reliability modelling, software testing methods and software efficiency. He is the author and co-author of many scientific publications. His paper appear in numerous journals, e.g. Control and Cybernetics, Polish Journal of Environmental Studies, Research in Logistics & Production and proceedings of international conferences, e.g. Depcos-Relcomex, Information Management or System, Modelling and Control.