Method of Increase of Check Efficiency Based on Small Size Samples

Dolgov A.Yu.
Department ITACMP
T.G. Shevchenko Pridnestrovien State University
Tiraspol, Republic of Moldova
dolgov@spsu.ru

***Abstract* – In presented article is offered one more decisive rule of selective check – the equivalent operational characteristic (EOC) which is based on the Weibull distribution. Settlement formulas are given and the table of an assessment of accuracy of calculation of predicted reject by the offered technique is given. The table of the comparative analysis of accuracy of all offered before methods and the existing standard is also submitted. It is proved that use of EOC based on the Weibull distribution gives more exact (by 1,05-5,30 times) results, than use of EOC based on other distribution laws, and 2,5-8,8 times more precisely, than operating methods of boundary check on samples of small size sample.**

***Key words* – sample check, the Weibull distribution, equivalent operational characteristic.**

I. Introduction

By production of crystals of integrated microcircuit (IMC) there was a practice when judge quality of 400-5000 working crystals by results of control measurements in 5 (or 10) the test cells (TC). Such ratio of objects of the control and certified samples isn't provided by any operational characteristics therefore on production were compelled to resort to an artificial method of borders that led to considerable false reject with big economic losses [1].

In the previous works [2-6] we found methods of essential increase of accuracy of forecasting of reject due to use of the method of pointed distributions (MPD) which is based on knowledge of the law of distribution of checked parameter, and the equivalent operational characteristic (EOC). As a result it is proved that the subjective component of false reject depending generally on an error of calculation of a mean square deviation (SD) decreased by 2,5-5,3 times that at introduction in production gives considerable economic effect. In article on the basis of development of MTR ways of further increase in accuracy of an assessment of possible reject in relation to each plate of party on the example of the most widespread law of distribution by production of crystals of IMC – Weibull's law which special cases an exponential and normal laws are considered.

II. Modifications of the Method of Pointed Distributions

Since the end of the 90th years and to the present by us it is offered and the method of pointed distributions (MPD) [2] which principle is based on rather wide circulation: each measurement should be considered as distribution center with the known law. It allows to pass from the volume of initial small sample of **n** to equivalent (virtual) sample which volume of $\mathbf{n_e}$ can be found through Kolmogorov's *D*-statistics and D_n-statistics similar to it considering that the amount of information in both cases is identical to the same confidential probability

$$n_e = n\left(\frac{D}{D_n}\right)^2 . \qquad (1)$$

An assessment of a relative error of one of the most sensitive statistical characteristics – *S* average quadratic deviation (AQD) – we will conduct on the formula offered by Y.B. Shor [8]

$$\varepsilon = \frac{\sigma(S)}{\sigma} \cdot 100\% = \sqrt{\frac{1}{2n - 1{,}4}} \cdot 100\% . \qquad (2)$$

Results of calculations for some volumes of initial sample and two types of laws of distribution at confidential probability P_{conf}=0,95 are presented in tables 1 and 2. From this tables it is visible that application of MPD allows to increase the accuracy (to reduce a mistake) of calculation by 1,7-2,0 times in case of bilateral distribution and by 2,3-2,5 times in case of unilateral distribution. Improvement of a method of pointed distributions led to creation of the robust method of pointed distributions (RMPD) [4] and at the same time combined method of pointed distributions (KMPD) [5].

The Robust method of pointed distributions differs from initial MPD in use of a steady (robust) assessment of Hodges-Lehman of an arithmetic average of sample on Walsh's averages. Thus the accuracy of calculations of selective AQD increases in average by 1,2-1,5 times in comparison with classical methods (see tab. 1 and 2).

Unfortunately, RMPD at calculation of the central moments of the second and the highest orders (and, so and AQD) had a systematic mistake because of a little reduced scope of equivalent sample. For elimination of this shortcoming, it is offered to add the volume of initial sample of ***n*** to the intermediate volume of equivalent robust sample of $\boldsymbol{n_e}$ of uniforms. Results of calculations are presented in tables 1 and from which analysis it is clear that the accuracy of calculations of selective AQD for KMPD increases in average by 1,5-2,0 times in comparison with MPD or by 2,9-3,4 times in comparison with classical methods of calculation.

TABLE 1. REAL **n** AND EQUIVALENT **ne** VOLUMES OF SAMPLES FOR DIFFERENT MODIFICATIONS OF MPD

Initial sample		3	4	5	6	7	8	9	10
Bilateral distributions	MPD	10	14	17	20	23	25	27	29
	PMPD	10	20	29	43	59	77	82	92
	KMPD	20	29	42	59	77	82	92	108
	OMPD	36	51	62	73	85	94	102	112
Unilateral distributions	MPD	13	23	33	34	37	40	41	42
	PMPD	13	34	42	55	66	79	83	96
	KMPD	33	42	55	66	79	83	96	108
	OMPD	49	80	97	108	119	135	145	154

TABLE 2. RELATIVE ERRORS OF SELECTIVE AQD FOR DIFFERENT MODIFICATIONS OF MPD, %

Sizes of initial samples, *n*		3	4	5	6	7	8	9	10
Bilateral distributions	Classic.	46,6	38,9	34,1	30,7	28,2	26,2	24,5	23,2
	MPD	23,2	19,4	17,5	16,1	15,0	14,3	13,8	13,3
	PMPD	23,2	16,1	13,3	10,9	9,3	8,1	7,8	7,4
	KMPD	16,1	13,3	11,0	9,3	8,1	7,8	7,4	6,8
	OMPD	11,8	9,9	9,0	8,3	7,7	7,3	7,0	6,7
Unilateral distributions	MPD	20,0	15,1	12,4	12,3	11,7	11,3	11,1	11,0
	PMPD	20,0	12,3	11,0	9,6	8,8	8,0	7,8	7,2
	KMPD	12,3	11,0	9,6	8,8	8,0	7,7	7,2	6,7
	OMPD	10,1	7,9	7,2	6,8	6,5	6,1	5,9	5,7

For understanding of further steps on increase in accuracy of calculations it was carried out graphic-analytical comparison of a classical method, MPD, KMPD on a concrete production example.

Let at the next operation of selective control in production of integrated chips crystals on one of plates sample with five test cells 9,1 was received; 10,2; 11,5; 12,8; 16,9 Ohms/sq, distributed in normal law. Then the interval estimates of AQD *S* received by various methods can be presented in the drawing form.

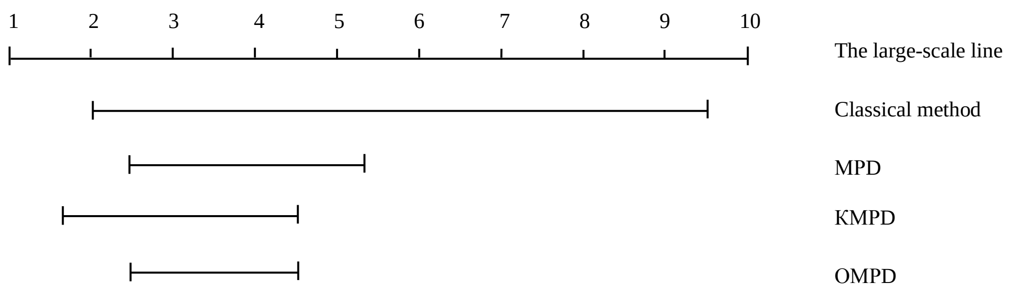


Fig. 1. Graphic interpretation of interval estimates of AQD by various methods.

As in all cases initial sample is the same, and the difference consists only in methods of calculation of its parameters thus the fundamental principle is observed: general parameter with the established confidential probability are contained between the lower and top borders of an interval assessment, further increase of accuracy of estimates is reached by use as a support of the left border of an interval assessment of MPD, and the right border – an interval assessment of KMPD for the same parameter arises. Then all estimates of a new method – the integrated method of pointed distributions (IMPD) – can be calculated as if back proceeding from borders of an interval assessment of AQD – left for MPD and right for KMPD. Results of calculations are presented in last lines of table 1 and 2.

The analysis of these results showed that OMPD allows to increase the accuracy of calculation of selective AQD (reduce a calculation error) in average by twice in comparison with MPD, by 3,5-4,0 times in comparison with classical methods for unimodal symmetrical distribution, by 1,3-2,0 times in comparison with MPD and by 4,0-4,6 times in comparison with classical methods for unilateral laws of distribution.

III. Increase of Accuracy of Estimates of Control Sample Parameters

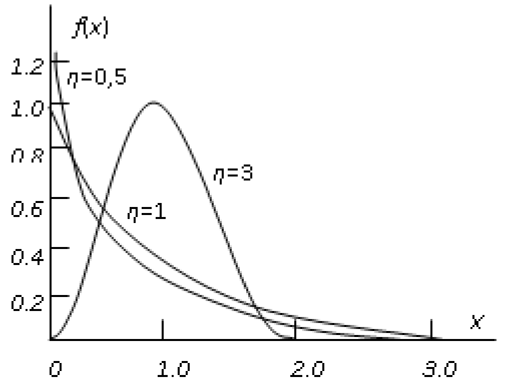
Function of Weibull distribution can be presented in the form [5]

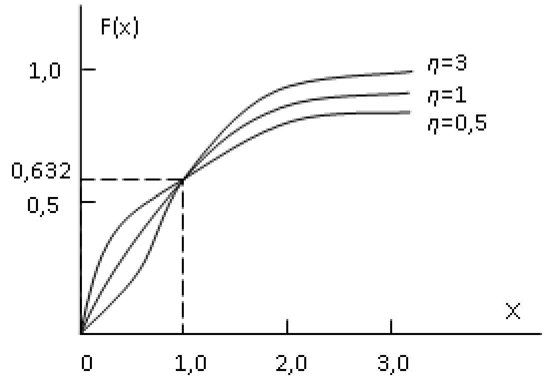
$$F(x) = 1 - exp\left[-\left(\frac{X-\theta}{b}\right)^{\eta}\right], \quad (3)$$

and its density (probability density) is

$$f(x) = \frac{\eta}{b}\left(\frac{X-\theta}{b}\right)^{\eta-1} exp\left[-\left(\frac{X-\theta}{b}\right)^{\eta}\right], \quad (4)$$

where b – scale parameter (sometimes $b = \frac{1}{\lambda}$); η – form parameter; θ – shift parameter (fig. 2).





1 2

Fig. 2. Density of probability (1) and function of Weibull distribution (2) at b=1, θ =0.

Calculation of parameters of distribution of Weibull represents a complex challenge [5], however at $\eta \geq 1$ it can be considerably simplified by means of the following approximations [6, c.151]:

$$\hat{\eta} = 4{,}8\,(r_3 + 1{,}23)^{-1{,}4};\ \hat{b} = \frac{\delta}{G};\ \theta = \overline{X} - \delta;$$
$$\delta \approx \left(0{,}5 + 0{,}784\hat{\eta} - \frac{0{,}35}{\hat{\eta}}\right) \cdot S;$$
$$G = \Gamma\left(1 + \frac{1}{\eta}\right) \approx 1 - 0{,}427(\eta - 1)\eta^{-1{,}9}, \quad (5)$$

where $\overline{X}$ – a sample average; S – AQD; r_3 – the third main issue.

If shift is absent (θ=0), estimates of parameters have an appearance

$$\hat{\eta} \approx \frac{n-1}{n}\left(0{,}465\frac{S}{\overline{X}} + 1{,}282\frac{\overline{X}}{S} - 0{,}7\right);\ \hat{b} = \frac{\overline{X}}{G}. \quad (6)$$

Separate samples by the volume of n=5 or n=10 (quantity of the test cells (TC) on each plate) can be checked for compliance to Weibull's distribution by means of Smirnov-Kramer-von Mises criterion

$$n\omega^2 = \frac{1}{12n} + \sum_{i=1}^{n}\left\{F\left(X_i - \frac{2i-1}{2n}\right)\right\}^2 \leq n\omega^2(P_{conf}), \quad (7)$$

where $F(X_i)$ – theoretical function of distribution.

It is necessary to remember that theoretical function of distribution (in our case Weybull distribution [7]) has to be known to within parameters. By researches [8] it is established that use of $F(X)$ as a function of distribution with the parameters estimated on sample (a widespread mistake!) leads to increase in quantity of errors of the second sort.

We looked for the solution of an objective by the way we have already found: equivalent (virtual) increase in volume of sample by means of the method of pointed distributions (MPD) [2] and application of the equivalent operational characteristic (EOC) [3].

To use a formula (3) or (4) for determination of parameters of Weibull distribution, it is necessary to find an arithmetic average of both the second central and the third main moments of each concrete sample (i.e. each concrete plate). From the point of view of MPD it means to find [9]

$$m_X^* = \frac{\sum_{j=1}^{k}\sum_{i=1}^{n} p_{ij}\cdot X'_j \cdot \exp\left[-4{,}5\left(\frac{X'_j - X_i}{\rho}\right)^2\right]}{\sum_{j=1}^{k}\sum_{i=1}^{n} p_{ij}\cdot \exp\left[-4{,}5\left(\frac{X'_j - X_i}{\rho}\right)^2\right]} \quad (8)$$

$$\mu_2^* = \frac{\sum_{j=1}^{k}\sum_{i=1}^{n} p_{ij}\cdot (X'_j)^2 \cdot \exp\left[-4{,}5\left(\frac{X'_j - X_i}{\rho}\right)^2\right]}{\sum_{j=1}^{k}\sum_{i=1}^{n} p_{ij}\cdot \exp\left[-4{,}5\left(\frac{X'_j - X_i}{\rho}\right)^2\right]} - (m_X^*)^2 \quad (9)$$

$$\mu_3^* = \frac{\sum_{j=1}^{k}\sum_{i=1}^{n} p_{ij}\cdot (X'_j)^2 \cdot \exp\left[-4{,}5\left(\frac{X'_j - X_i}{\rho}\right)^2\right]}{\sum_{j=1}^{k}\sum_{i=1}^{n} p_{ij}\cdot \exp\left[-4{,}5\left(\frac{X'_j - X_i}{\rho}\right)^2\right]} - 3\mu_2^* m_X^* - (m_X^*)^3 \quad (10)$$

$$r_3^* = \frac{\mu_3^*}{(\sqrt{\mu_2^*})^3}. \quad (11)$$

Equivalent size of samples for Weibull distribution can be found from expression (10)

$$n_e = \sqrt{-713{,}7 + 301{,}2n - 6{,}907n}, \quad (12)$$

that means for n=5 → n_e=25; for n=10 → n_e=40.

Then equivalent operational characteristic is

$$P(q) = F_\circ\left(\left(U_{1-q} - \frac{k_s}{k_n}\right)\Big/\sqrt{\frac{1}{n} + \frac{k_s^2}{2n-1{,}4}}\right),$$

where $k_n = \sqrt{\frac{n-1}{2}}\cdot\Gamma\left(\frac{n-1}{2}\right)/\Gamma\left(\frac{n}{2}\right)$ – correction index, and $F_\circ(\cdot)$ – Gauss's integral, at substitution of volume **n** of initial sample by volume $\mathbf{n_e}$ of equivalent samples. Taking into account concrete values of admissible errors of the first sort (for example, α=0,10), and also threshold of 100% acceptance q_0 (for example, q_0=0,10), we will receive

for n=5 and n_e=25 value k_n=k_{25}=0,10105; k_s=1,036;

for n=10 and $n_э$=40 value k_n=k_{40}=0,10064; k_s=1,107.

Then equivalent operational characteristics will be:

for n=5

$$P(q) = F_\circ\left(\frac{U_{1-q} - 1{,}025}{0{,}2492}\right) = \frac{|T - m_X^*|}{\sqrt{\mu_2^*}}; \quad (13)$$

for n=10

$$P(q) = F_\circ\left(\frac{U_{1-q} - 1{,}100}{0{,}2015}\right) = \frac{|T - m_X^*|}{\sqrt{\mu_2^*}}, \quad (14)$$

where T – the upper or lower bound of norm, or, taking into account the law of distribution, will look as

$$P(q) = \exp\left[-\left(\frac{X-\theta}{b}\right)^\eta\right] \quad (15)$$

We will remind that range of the predicted reject has dual origin from the objective and subjective reasons. The objective reasons are the following: the volume of sample n, coefficient of overlapping of norm $\upsilon = (X_{max} - X_{min})/(T_B - T_H)$ and arrangement of concrete values of controlled parameter on a numerical axis (see fig. 3), to the subjective reasons refers a big mistake of calculation of a mean square deviation (table 3).

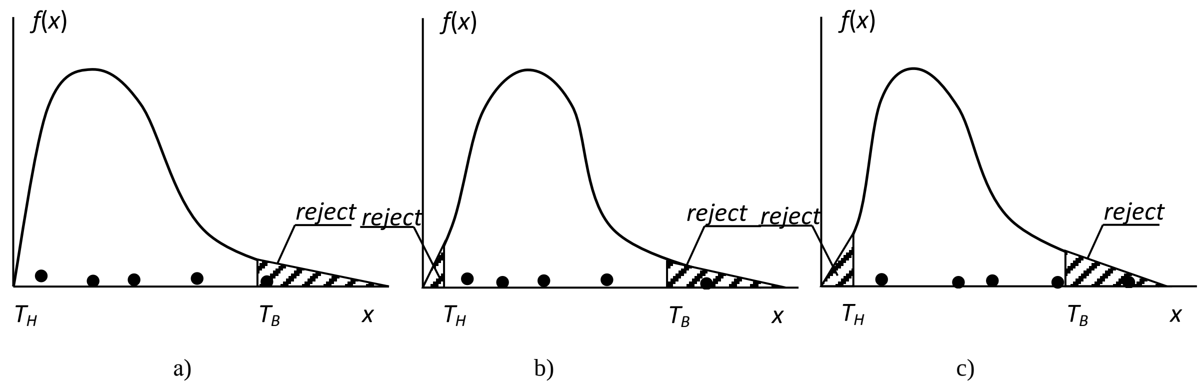


Fig. 3. Options of an arrangement of measurements at m=4 and υ=1.
a) minimum reject; b) the most probable reject; c) maximum reject

TABLE 3. SIZES OF THE PREDICTED REJECT (%) ON A PLATE WITH FIVE TEST CELLS (NE=25)

The predicted reject	$\sqrt{\mu_2^*}$	Number of the measurements which aren't going beyond norm, *m*										
		m=5			m=4				m=3			
		υ=1,0	*υ*=0,7	*υ*=0,5	*υ*=1,3	*υ*=1,0	*υ*=0,7	*υ*=0,5	*υ*=2,0	*υ*=1,0	*υ*=0,7	*υ*=0,5
Minimum	**Min.**	–	0,6	0,1	4,7	2,2	2,8	3,3	13,3	5,9	4,9	4,9
	Aver.	–	3,6	0,6	12,3	8,2	14,9	15,9	25,6	19,9	19,8	19,5
	Max.	–	6,5	2,6	19,4	13,3	14,9	16,1	32,5	21,6	19,9	19,6
The most probable	**Min.**	–	1,0	0,7	4,9	3,5	3,4	4,0	14,3	7,9	6,0	5,9
	Aver.	–	7,1	6,9	15,2	14,3	16,4	17,8	26,9	23,1	21,4	21,4
	Max	–	8,8	7,3	19,8	16,6	16,4	17,8	33,8	25,1	21,8	21,6
Maximum	**Min.**	2,2	1,9	2,3	6,6	6,1	3,5	4,7	14,7	9,4	8,0	6,4
	Aver.	8,0	11,9	13,5	19,7	20,6	16,6	19,3	27,8	25,3	24,7	22,2
	Max.	13,1	12,2	13,6	22,9	21,9	16,6	19,3	34,2	27,3	25,3	22,5

The similar table can be provided for a plate case with ten test cells.

For evaluation of reject forecasting accuracy (or decrease in a share of subjective factors in forecasting), it is necessary to compare results of calculations for different methods. For this purpose in table 4 are presented the relation of the predicted reject by three techniques to the predicted reject on EOC taking into account Weibull distribution.

TABLE 4. RELATIVE REDUCTION OF THE MAXIMUM REJECT (TIMES)

The predicted reject	AQD	Number of the measurements which aren't going beyond norm, *m*										
		m=5			m=4				m=3			
		υ=1,0	*υ*=0,7	*υ*=0,5	*υ*=1,3	*υ*=1,0	*υ*=0,7	*υ*=0,5	*υ*=2,0	*υ*=1,0	*υ*=0,7	*υ*=0,5
According to the existing standards	**Min.**	–	8,83	16,58	3,96	5,22	3,89	2,94	2,56	3,22	3,01	2,58
	Aver.	–	6,56	5,09	3,88	4,20	3,59	2,75	2,50	2,81	2,82	2,46
	Max.	5,30	4,75	3,42	3,35	3,20	3,60	2,62	2,50	2,62	2,49	2,48
On EOC of the normal law	**Min.**	–	5,28	12,81	1,84	2,65	2,30	2,07	1,11	1,63	1,73	1,70
	Aver.	–	3,90	4,56	1,80	2,13	2,09	1,88	1,07	1,41	1,58	1,55
	Max..	2,68	2,81	2,46	1,55	1,63	2,07	1,73	1,05	1,30	1,36	1,49
On EOC of the exponential law	**Min.**	–	1,26	1,27	1,25	1,26	1,27	1,25	1,25	1,25	1,25	1,25
	Aver.	–	1,25	1,25	1,25	1,25	1,26	1,25	1,25	1,25	1,25	1,25
	Max.	1,25	1,25	1,25	1,25	1,25	1,25	1,25	1,25	1,25	1,25	1,25

IV. CONCLUSION

The analysis of table 4 showed that application of the equivalent operational characteristic on base of Weibull distribution law yields the results by 2,5-8,8 times more exact, than the existing standards, by 1,05-5,30 times more exact, than EOC with normal distribution law, by 1,25 times more exact, than EOC with exponential distribution (sharply different results: 16,8 in the first case and 12,81 – in the second, didn't take into account as casual deviations). Thus, all offered methods of assessment of results of control on small size samples based on equivalent operational characteristics taking into account laws of distribution (especially Weibull distribution), can be recommended for implementation in the production.

REFERENCES

[1] A.Y. Dolgov, "Increase of efficiency of statistical control methods and management of technological processes of production of chips"// Dis …. Cand. Tech. Sci.: 05.11.13/MGAPI. – M, 2000. – 217 p.

[2] A.Y. Dolgov, Y.A. Dolgov and Y.A. Stolyarenko, "Method of increase of accuracy of calculation of parameters of sample of small volume (method of pointed distributions)"// Vestn.of Pridnestr. state. university. – 2010. – Anniversary edition – pp. 232-242.

[3] A.Y. Dolgov, "Increase of accuracy of estimates of parameters of control sample of small volume" // Radio-electronics and computer systems. – 2012. – No. 6 (58). – Kharkiv: KAI, 2012. – pp. 119-123.

[4] A.Y. Dolgov, "Increase of accuracy of estimates of parameters of control sample of small volume at the exponential law of distribution" // Radio-electronics and computer systems. – 2013. – No. 5(64). – Kharkiv: KAI, 2013. – pp. 135-138.

[5] A.Y. Dolgov, "Increase of accuracy of estimates of parameters of control sample of small volume at Weibull's distribution" // Radio-electronics and computer systems. – 2014. – No. 5(69). – Kharkiv: KAI, 2014. – pp. 110-114.

[6] A.S. Grunichev, A.I. Mikhaylov, Y.B. Shor, "Tables for calculation of reliability at Weibull's distribution" – M.: Publishing house of standards, 1974. – 64 p.

[7] A.I. Kobzar, "Applied mathematical statistics. For engineers and scientists" – M.: FIZMATLIT, 2006. – 816 p.

[8] Y.B. Shor and F.I. Kuzmin, "Statistical methods of the analysis and quality control and reliability" – M.: Sov. radio, 1968. – 213 p.

[9] A.I. Orlov, "A widespread mistake when using criteria of Kolmogorov and an omega square"// Plant. lab. – 1985. – t.51. – No. 1. – pp. 60-62.

[10] Y.A. Dolgov, "Research of interval estimates and the moments of the highest order for samples of small volume"// Radio-electronics and computer systems. – 2012. – No. 5 (57). – Kharkiv: KAI, 2012. – pp. 165 – 170.

[11] Y.A. Dolgov, "Scoping of equivalent sample in a method of pointed distributions"// Radio-electronics and computer systems. – 2013. – No. 5 (64). – Kharkiv: KAI, 2013. – pp. 139-141.