UNIVERSITY OF PITESTI SCIENTIFIC BULLETIN

Faculty Of Mechanics And Technology AUTOMOTIVE series, year XXI, no. 25

CONTRIBUTIONS TO THE ANALYSIS OF MEASUREMENT DATA OBTAINED FOR ATTRIBUTE TYPE CHARACTERISTICS

Alin Rizea
University of Pitesti, Romania

Article history:
Received: 22.01.2016; Accepted: 18.03.2016.

Abstract:
The control process can be affected by a series of errors. In order to obtain the results wanted, these errors have to be removed (or diminished). Some of these errors are due to both the operator and the means by which the control is accomplished. In order to diminish their effect there can be used a method to study the capability of the equipment (method) used for the measurement. The R&R study can be carried out both before starting the measurement process and during the measurement.

Keywords: Repeatability & Reproducibility, Hypothesis Test Analyses, Cross-Tab Method

INTRODUCTION

The decision to adjust the operating parameters of a production process is frequently based on the analysis of measurement data. The measurement data or some statistical data computed are compared to the statistical control limits of the process and if following the comparison it results that the process is outside these established limits, then, there will be made a certain type of adjustment. Otherwise, the process will be carried out without any adjustment.
Measurement errors are errors which appear during the measurement process. During this process, due to the action of disruptive factors (imperfection of measurement means, variation of the conditions under which the measurement is carried out, variation of the sizes measured, qualification and attention of the operator who carries out the measurement etc.) the values of the sizes measured cannot be determined with absolute precision. They are affected by measurement errors.
By measurement error one understands the difference between the result of the measurement of a given size (x) and the real value of a measured size (x_0); meaning:

$$E = x - x_0 \tag{1}$$

Of course, usually, the real value of the size is unknown. Often, one has an idea on what might be the real value of a size, from previous experiments (including other measurement means) or from a theoretical approach. These previous pieces of information help us assess the order of size of the value we are expecting after a measurement. It is desirable to find, using the experimental data, a procedure to determine how much confidence we can have in these.
For a certain measurement method, the total measurement error (E_T) is a statistical term formed by the sum of errors composing it.

PRINCIPLE OF AN R&R STUDY

In industry the measurement systems are generally equipment which measures diameters, weights, lengths or other physical sizes.

In case the quality characteristics monitored are attribute type size, the products are classified as “good” or “with flaws” (acceptable units or unacceptable units) or are ranked according to the number of flaws detected (number of unacceptable characteristics). The attribute type data generally have only two values: acceptable/unacceptable, accepted/rejected, pass/no pass, present/absent. And the measurable sizes can be transformed into attribute type measures: for example, the concordance of the diameter of a reaming, when it is measured with the help of a pass/no pass calibre.
An attribute tool is the one which compares each piece of a specific set and accepts the piece if the limits are satisfied, otherwise it rejects the piece. Most of these tools are adjusted to accept or reject a set of sample pieces.
Unlike a variable tool, an attribute tool can only indicate how good or bad a piece is and not the fact that this piece is accepted or rejected.
In order to control the subjectivity of these quality measurements, the same characteristics will be measured by several persons and compared. If the conclusions drawn are coherent, they are probably valid, but if they do not match they have limited use.
During an R & R study (Répétabilité & Reproductibilité – in French; Repeatability & Reproducibility – in English) the capability of the control means used are being checked.
The measurement error is mainly determined by two factors [1]:
- The measurement equipment which makes for a set of measurements carried out under the same conditions to exist a certain variation characterised by “repeatability”;
- The operator who carries out the measurements. It is obvious that the values obtained through measurement by several operators under the same conditions will be different, this variation is characterised by the term of “reproducibility” [2].
Usually, an R&R study is carried out before starting the measurement process and during the measurement. At the beginning, the capability of the measurement equipment (method) used is checked. Along the way, it is repeated each time a factor (equipment, operator) is changed and an annual check is carried out in order to control the variation tendency of the measurement process[3].

PREPARING THE SAMPLE

The attribute tool selected for the study is a double calibre type PASS/no PASS (fig. 1).

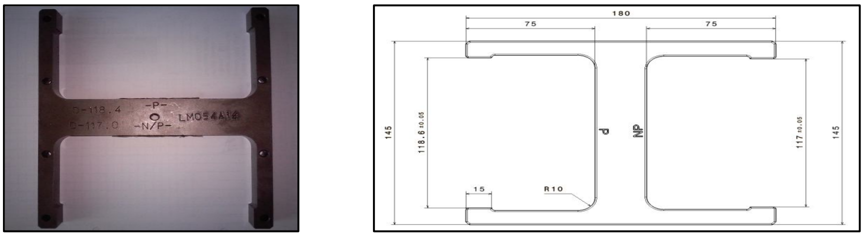

Figure 1. Calibre PASS/no PASS

This calibre will be used to check the outer diameter of a pipe type part (fig. 2).

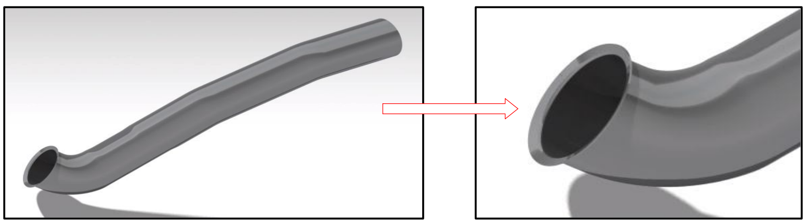

Figure 2. Diameter to check

COLLECTING DATA

The samples used for the study (30 pieces) are not randomly chosen. They are chosen by specialised personnel and have to be determined as being acceptable or unacceptable (their status is known by only one person). The acceptable parts were noted with "OK" and the unacceptable ones with "NOK". Parts whose state of reference is given in table 1 are measured 3 times by 3 different assessors (independently from each other). The results of the measurements are also presented in table 1.

Table 1. Results of the checks

		Evaluator 1			Evaluator 2			Evaluator 3		
Referință	Nr. Eșantion	1A	2A	3A	1B	2B	3B	1C	2C	3C
OK	1	OK	OK	OK	OK	OK	OK	OK	OK	OK
OK	2	OK	OK	OK	OK	OK	OK	OK	OK	OK
NOK	3	NOK	NOK	NOK	NOK	OK	OK	NOK	OK	NOK
NOK	4	NOK	NOK	NOK	NOK	NOK	NOK	NOK	NOK	NOK
NOK	5	NOK	NOK	NOK	NOK	NOK	NOK	NOK	NOK	NOK
OK	6	OK	OK	OK	OK	OK	OK	OK	OK	OK
OK	7	OK	OK	OK	OK	OK	OK	OK	OK	OK
OK	8	OK	OK	OK	OK	OK	OK	OK	OK	OK
NOK	9	NOK	NOK	NOK	NOK	NOK	NOK	NOK	NOK	NOK
OK	10	OK	OK	OK	OK	OK	OK	OK	OK	OK
OK	11	OK	OK	OK	OK	NOK	OK	OK	NOK	OK
NOK	12	NOK	NOK	NOK	NOK	OK	NOK	NOK	NOK	NOK
OK	13	OK	NOK	OK	OK	OK	OK	OK	OK	OK
OK	14	OK	OK	OK	OK	OK	OK	OK	OK	OK
OK	15	OK	OK	OK	OK	OK	OK	OK	OK	OK
OK	16	OK	OK	OK	OK	OK	OK	OK	OK	OK
OK	17	OK	OK	OK	OK	OK	OK	OK	OK	OK
OK	18	OK	OK	OK	OK	OK	OK	OK	OK	OK
OK	19	OK	OK	OK	OK	OK	OK	OK	OK	OK
OK	20	OK	OK	OK	OK	OK	OK	OK	OK	OK
OK	21	NOK	OK	NOK	OK	OK	OK	OK	OK	OK
NOK	22	OK	OK	OK	NOK	NOK	NOK	NOK	NOK	NOK
OK	23	OK	OK	OK	OK	OK	OK	OK	OK	OK
OK	24	OK	OK	OK	OK	OK	OK	OK	OK	OK
NOK	25	NOK	NOK	NOK	NOK	NOK	NOK	NOK	NOK	NOK
NOK	26	NOK	NOK	NOK	NOK	OK	NOK	NOK	NOK	NOK
OK	27	OK	OK	OK	OK	OK	OK	OK	OK	OK
OK	28	OK	OK	OK	OK	OK	OK	OK	OK	OK
OK	29	OK	OK	OK	OK	OK	OK	OK	OK	OK
NOK	30	NOK	NOK	NOK	NOK	NOK	NOK	NOK	NOK	NOK

In the case of using attribute type tools there is the risk to obtain contradictory information about a certain measurement. In order to analyse this risk there can be used the Hypothesis Test Analyses. The use of this test can determine the degree of agreement between individual assessments, but also between individual and reference assessments.

HYPOTHESIS TEST ANALYSIS - CROSS-TAB METHOD

Based on table 1 the data were expressed in pairs of assessors considering the number of times the two assessors agreed or did not agree for each set of assessments (the results of comparisons between the assessments carried out by Assessor 1 and Assessor 2 are given as examples in table 2). The distribution of the data waited was estimated as well. Using the relations below, there was computed the probability that a pair of assessors may/may not agree on a measurement. Considering that the two assessors are independent the probability that both of them agree that a piece is NOK is given by the product of the two probabilities:

$$P(P_{Ev1NOK} \cap P_{Ev2NOK}) = P_{Ev1NOK} \text{ x } P_{Ev2NOK} = (25/90) \text{ x } (28/90) \tag{2}$$

The number of expectations each time Assessor 1 and Assessor 2 agree that the piece is NOK is estimated by multiplying the probability combined to the number of observations:

$$90 \text{ x } (P_{Ev1NOK} \text{ x } P_{Ev2NOK}) = 90 \text{ x } (25/90) \text{ x } (28/90) = 7.78 \tag{3}$$

The number of expectations each time Assessor 1 and Assessor 2 agree that the piece is OK is computed similarly:

$$P(P_{Ev1OK} \cap P_{Ev2OK}) = P_{Ev1OK} \text{ x } P_{Ev2OK} = (65/90) \text{ x } (62/90) \quad (4)$$

$$90 \text{ x } (P_{Ev1OK} \text{ x } P_{Ev2OK}) = 90 \text{ x } (65/90) \text{ x } (62/90) = 44.78 \quad (5)$$

The values thus obtained are presented in table 2.

Table 2

			Assessor 2		
			NOK	OK	Total
Assessor 1	NOK	Numbered	22	6	28
		Waited	7.78	20.22	28
	OK	Numbered	3	59	62
		Waited	17.22	44.78	62
	Total	Numbered	25	65	90
		Waited	25	65	90

		Assessor 2				
		NOK	OK	Total	% NOK	% OK
Assessor 1	NOK	22	6	28	79	21
	OK	3	59	62	5	95
	Total	25	65	90		

Similarly there have been computed for the other pairs of assessors as well (assessor 1 with assessor 3 and assessor 2 with assessor 3). The results are presented in table 3 and table 4:

Table 3

			Assessor 1		
			NOK	OK	Total
Assessor 3	NOK	Numbered	25	2	27
		Waited	8.4	18.6	27
	OK	Numbered	3	60	63
		Waited	19.6	43.4	63
	Total	Numbered	28	62	90
		Waited	28	62	90

		Assessor 1				
		NOK	OK	Total	% NOK	% OK
Assessor 3	NOK	25	2	27	93	7
	OK	3	60	63	5	95
	Total	28	62	90		

Table 4

			Assessor 2		
			NOK	OK	Total
Assessor 3	NOK	Numbered	24	3	27
		Waited	7.5	19.5	27
	OK	Numbered	1	62	63
		Waited	17.5	45.5	63
	Total	Numbered	25	65	90
		Waited	25	65	90

		Assessor 2				
		NOK	OK	Total	% NOK	% OK
Assessor 3	NOK	24	3	27	89	11
	OK	1	62	63	2	98
	Total	25	65	90		

Next there will be determined the degree of agreement between the assessors. In order to do this determination there was used the *k test (kappa test)* which measures the agreement between the assessments of 2 assessors when they verify the same piece. If the two totally agree, then k = 1, and if they totally disagree (agree only due to chance), then k ≤ 0 [4, 6]. Coefficient k is computed with the help of formula:

$$\text{Kappa} = \frac{p_0 - p_e}{1 - p_e}, \tag{6}$$

where p_o represents the relative agreement between the assessors and is computed with the help of formula:

$$p_0 = \frac{AgOK + AgNOK}{Tot\ (OK + NOK)} \tag{7}$$

p_e represents the probability to obtain an agreement between the assessors and is computed with the help of formula:

$$p_e = \frac{Waited\ OK + Waited\ NOK}{Tot\ (OK + NOK)} \tag{8}$$

By using these formulas there was computed the coefficient k for each pair of assessors. For Assessor 1 and Assessor 2 we have:

$$p_o = \frac{22+59}{90} = 0.90; \qquad p_e = \frac{7.78+44.78}{90} = 0.58 \text{ kappa} = \frac{0.90-0.58}{1-0.58} = 0.76 \tag{9}$$

Similarly there was computed for the other pairs of assessors and the results are presented in table 5.

Table 5. Degree of agreement between assessors

	Assessor 1	Assessor 2	Assessor 3
Assessor 1	-	0.760	0.817
Assessor 2	0.760	-	0.892
Assessor 3	0.817	0.892	-

In order to interpret the results obtained there will be used the following table (proposed by Landis and Coch in paper [3]).

Table 6

Value of coefficient k	< 0	0.0 – 0.20	0.21 – 0.40	0.41 – 0.60	0.61 – 0.80	0.81 – 1.0
interpretation	disagreement	Very weak agreement	Weak agreement	Moderate agreement	Very good agreement (strong)	Almost perfect agreement

This analysis shows the fact there is a very good agreement between assessors, but this does not inform us on how well the measurement system sorts the good pieces from the bad ones.
For the following analysis the pieces were assessed with the help of variable measurement system by the quality inspector in order to determine the reference values. In order to obtain this the results of each assessor were compared to the reference values. There were used the same methods and the data were centralized in the following tables:

Table 7

			Reference		
			NOK	OK	Total
Assessor 1	NOK	Numbered	26	2	28
		Waited	8.4	19.6	28
	OK	Numbered	1	61	62
		Waited	18.6	43.4	62
	Total	Numbered	27	63	90
		Waited	27	63	90

		Reference				
		NOK	OK	Total	% NOK	% OK
Assessor 1	NOK	26	2	28	93	7
	OK	1	61	62	2	98
	Total	27	63	90		

Table 8

			Reference		
			NOK	OK	Total
Assessor 2	NOK	Numbered	23	2	25
		Waited	7.5	17.5	25
	OK	Numbered	4	61	65
		Waited	19.5	45.5	65
	Total	Numbered	27	63	90
		Waited	27	63	90

		Reference				
		NOK	OK	Total	% NOK	% OK
Assessor 2	NOK	23	2	25	92	8
	OK	4	61	65	6	94
	Total	27	63	90		

Table 9

			Reference		
			NOK	OK	Total
Assessor 3	NOK	Numbered	26	1	27
		Waited	8.1	18.9	27
	OK	Numbered	1	62	63
		Waited	18.9	44.1	63
	Total	Numbered	27	63	90
		Waited	27	63	90

		Reference				
		NOK	OK	Total	% NOK	% OK
Assessor 3	NOK	26	1	27	96	4
	OK	1	62	63	2	98
	Total	27	63	90		

In this case as well the kappa measure was computed in order to determine the agreement of each assessor to the reference values (table 10).

Table 10

	Assessor 1	Assessor 2	Assessor 3
kappa	0.87	0.84	0.95

These values show the fact that there is a good agreement between the assessors and the reference.

CONCLUSIONS

Considering the information presented in the article the following conclusions can be drawn:

- In the case of using attribute type measurement tools there is the risk to obtain contradictory information about a measurement;
- With the help of the method presented in the article, this risk can be assessed and there can be determined the degree of agreement both between individual assessment and between individual and reference assessments;
- As it can be noticed from the data presented, there is a very good concordance both between the assessors and between the results obtained by them and the reference results.

REFERENCES

1. Boroiu Al., *Analiza sistemelor de masurare*, Mentenanta si distributie industriala, Nr. 10, 2012
2. Boroiu Al., *Instrumente statistice utilizate in managementul calitatii*, Ed UPit, 2010
3. Chrysler, Ford, GM, *Measurement systems analysis,* AIAG, Detroit, Michigan. (2010)
4. Landis, J. R. and Koch, G. G., *The Measurement of Observer Agreement for Categorical Data*, Biometrics, Vol. 33, 1977, pp. 159—174
5. *ASTME 177-86, Use of the Terms Precision and Bias in ASTM Test Methods*
6. http://www.rubymetrology.com/add_help_doc/MSA_Reference_Manual_4th_Edition.pdf