Quality Engineering, 10(2), 263-266 (1997-98)

ON THE ESTIMATION OF GAUGE CAPABILITY

M. V. Asokan
Statistical Quality Control & Operations Research Unit
Indian Statistical Institute
1095 Mettupalayam Road
Coimbatore 641 002, India

Key Words

Repeatability; Reproducibility; Measurement error.

Introduction

Measurement error plays a significant role in the estimation of machine and process capability. If this error is significant, one may unnecessarily suspect the capability of the manufacturing process. The estimation of gauge capability is discussed in detail by Montgomery (1) and Montgomery and Runger (2).

Bias in the Estimator

In the classical method discussed by Montgomery and Runger (2), the estimates of gauge repeatability and reproducibility (GRR) are

$$\sigma_{(\text{repeatability})} = \frac{\overline{R}}{d_2}, \qquad \sigma_{(\text{reproducibility})} = \frac{R_x}{d_2}, \tag{1}$$

where $\overline{R}$ is the average range of the measurements done by each inspector on each part, R_x is the range of the average values of the measurements done by all inspectors, and d_2 is the scale factor to estimate the standard deviation from the range.

The scale factor d_2 depends on the size of the sample (n) under the assumption of an infinite (or large) number of samples (k). For the repeatability estimate, the number of samples (k) will be equal to the product of the number of parts and the number of inspectors, and the sample size (n) will be the number of measurements taken by each inspector on each part. In the case of the reproducibility estimate, the number of samples (k) will be one and the sample size (n) will be the number of inspectors. Montgomery and Runger (2) assumed the number of samples (k) infinite in the classical method. Because the number of samples (k) cannot be assumed to be infinite, the estimates in Eq. (1) are often biased. In the case of the reproducibility estimate discussed by Montgomery and Runger (2), the number of samples is only one and sample size is three.

Unbiased Estimators

Suppose we have k samples each of size n from a normal population with standard deviation σ. Let $R_1, R_2, \ldots, R_k$ be the sample ranges and $\overline{R}$ their average. Patnaik (see Ref. 3) showed that $\nu(\overline{R}/d_2^*)^2/\sigma^2$ is approximately distributed as chi-square with ν degrees of freedom, where the scale factor d_2^* and equivalent degrees of freedom ν are given as functions of n and k. In this case, the unbiased estimators of GRR are

$$\sigma_{(\text{repeatability})} = \frac{\overline{R}}{d_2^*},$$
$$\sigma_{(\text{reproducibility})} = \frac{R_x}{d_2^*}. \qquad (2)$$

The values of the scale factor d_2^*, together with the degrees of freedom ν, are given in Table 1, where n varies from 2 to 5 and k varies from 1 to 15. From Table 1, it can be seen that $d_2^* \geq d_2$, which results in the overestimation of GRR values if d_2 is used instead of d_2^*.

The percentage overestimation of GRR values for different values of n and k when d_2 is used instead of d_2^* is estimated and given in Table 2.

It can be seen from Table 2 that the percentage overestimation is highly significant for small values of n and k. Generally, for any gauge capability study, the maximum number of inspectors will be 3, the number of parts will be 10, and the number of measurements per part by inspector will be 2. The largest percentage overestimation of gauge reproducibility will be 24.8 if the number of inspectors involved in the study is 2 and is reduced to 13.0 for 3 inspectors. When the number of parts is 10, the percentage overestimation in gauge repeatability will be 1.4 for 2 inspectors and is reduced to 0.68 for 3 inspectors; that is, the reproducibility estimate is more overestimated than the repeatability estimate when d_2 is used instead of d_2^*. An overestimation of GRR values may unnecessarily call into question the capability of the manufacturing process.

Confidence Interval

Because $\nu(\overline{R}/d_2^*)^2/\sigma^2$ follows a chi-square distribution if the variable is normally distributed, the confidence interval for gauge repeatability and reproducibility can be easily estimated.

A $(100 - \alpha)\%$ confidence interval for σ^2 is

$$\frac{\nu s^2}{\chi^2_{\alpha/2,\nu}} \leq \sigma^2 \leq \frac{\nu s^2}{\chi^2_{(1-\alpha/2),\nu}},$$

Table 1. Values Associated with the Distribution of the Average Range

NO. OF SAMPLES	SIZE OF THE SAMPLE (n)							
	2		3		4		5	
(k)	ν[a]	d_2^*	ν	d_2^*	ν	d_2^*	ν	d_2^*
1	1.0	1.41	2.0	1.91	2.9	2.24	3.8	2.48
2	1.9	1.28	3.8	1.81	5.7	2.15	7.5	2.40
3	2.8	1.23	5.7	1.77	8.4	2.12	11.1	2.38
4	3.7	1.21	7.5	1.75	11.2	2.11	14.7	2.37
5	4.6	1.19	9.3	1.74	13.9	2.10	18.4	2.36
6	5.5	1.18	11.1	1.73	16.6	2.09	22.0	2.35
7	6.4	1.17	12.9	1.73	19.4	2.09	25.6	2.35
8	7.2	1.17	14.8	1.72	22.1	2.08	29.3	2.35
9	8.1	1.16	16.6	1.72	24.8	2.08	32.9	2.34
10	9.0	1.16	18.4	1.72	27.6	2.08	36.5	2.34
11	9.9	1.16	20.2	1.71	30.3	2.08	40.1	2.34
12	10.8	1.15	22.0	1.71	33.0	2.07	43.7	2.34
13	11.6	1.15	23.9	1.71	35.7	2.07	47.4	2.34
14	12.5	1.15	25.7	1.71	38.5	2.07	51.0	2.34
15	13.4	1.15	27.5	1.71	41.2	2.07	54.6	2.34
d_2		1.13		1.69		2.06		2.33
c.d.[b]	0.88		1.82		2.74		3.62	

[a] ν = degrees of freedom.
[b] c.d. = constant difference.
Source: From Ref. 3.

Table 2. Percentage Overestimation of GRR Values

	n			
k	2	3	4	5
1	24.8	13.0	8.7	6.4
2	13.3	7.1	4.4	3.0
3	8.8	4.7	2.9	2.1
4	7.1	3.6	2.4	1.7
5	5.3	3.0	1.9	1.3
6	4.4	2.4	1.5	0.9
7	3.5	2.4	1.5	0.9
8	3.5	1.8	1.0	0.4
9	2.7	1.8	1.0	0.4
10	2.7	1.8	1.0	0.4
11	1.8	1.2	1.0	0.4
12	1.8	1.2	0.5	0.4
13	1.8	1.2	0.5	0.4
14	1.8	1.2	0.5	0.4
15	1.8	1.2	0.5	0.4

where v is the degrees of freedom, s^2 is the standard deviation estimated by $\bar{R}/d_2^*$, and α is the confidence coefficient.

Case Study

The focus of all modern quality management systems is customer satisfaction. To achieve this, all the measurable quality characteristics should be aimed at the target with minimum variation. Thus, the quantification of measurement error is necessary to ensure a reliable control system.

This study was conducted at an ISO 9001 certified company which is engaged in the manufacture of pumps and valves. Although most of the components have an open tolerance, critical components have a tight tolerance which is very often difficult to meet. It is, therefore, necessary to estimate measurement error and to see the scope for reducing it to meet the tolerances.

A training program on statistical techniques was conducted for the inspectors, in which measurement system variation was covered. To gain "hands-on" experience, some groups were assigned GRR studies. Estimation procedures together with comparison of estimates using d_2 and d_2^* were discussed in detail for a particular component.

The width of a particular component supplied by the vendors is a critical quality characteristic. The width specification was 69 ± 0.4 mm. Two inspectors were chosen from the Goods Inward Inspection department and seven components were taken at random for the study. Both inspectors measured the width of all the components twice, using a dial vernier accurate to within 0.02 mm. The data on the gauge capability study are given in Table 3.

Estimation of Gauge Repeatability

Number of samples (k) = (Number of parts) (Number of inspectors) = 7(2) = 14

Samples size (n) = Number of measurements/part/inspector = 2

Average range of all components for both inspectors ($\bar{R}$) = 1.3/14 = 0.0929

	Using d_2	*Using d_2^**
Gauge repeatability (6σ)	$6(\bar{R}/d_2)$ 6(0.0929/1.13) 0.4933	$6(\bar{R}/d_2^*)$ 6(0.0929/1.15) 0.4847
Percentage over tolerance	61.7	59.6
Percentage overestimation	1.8	

Estimation of Gauge Reproducibility

Number of samples (k) = 1

Samples size (n) = Number of inspectors = 2

Average width for Inspector I = 69.6100

Average width for Inspector II = 69.6543

Range of inspector averages (R_x) = 0.0443

	Using d_2	*Using d_2^**
Gauge reproducibility (6σ)	$6(R_x/d_2)$ 6(0.0443/1.13) 0.2352	$6(R_x/d_2^*)$ 6(0.0443/1.41) 0.1885
Percentage over tolerance	29.4	23.6
Percentage overestimation	24.8	

From the study, it was found that the percentage overestimation was very high for the reproducibility estimate, and the repeatability estimate was very high compared to the reproducibility estimate. The major cause for the high value of the repeatability estimate was due to the mix-up of product variation with repeatability. This was because the in-

Table 3. Gauge Capability Study Data

COMPONENT NO.	INSPECTOR I			INSPECTOR II		
	1	2	RANGE	1	2	RANGE
1	69.38	69.60	0.22	69.62	69.52	0.10
2	69.72	69.80	0.08	69.78	69.90	0.12
3	69.58	69.70	0.12	69.70	69.62	0.08
4	69.50	69.50	0.00	69.46	69.50	0.04
5	69.48	69.40	0.08	69.50	69.42	0.08
6	69.56	69.40	0.16	69.68	69.64	0.04
7	69.90	70.02	0.12	69.94	69.88	0.06

spectors did not repeat their measurements at the same point. It is, therefore, indicated clearly that the GRR estimates are inclusive of some variation of the product itself. The entire study was repeated by measuring the width at the specific point marked and the GRR estimates arrived at using d_2^* are as follows:

Gauge repeatability (6σ)	= 0.1831
Percentage over tolerance	= 22.9
Gauge reproducibility (6σ)	= 0.0898
Percentage over tolerance	= 11.2
Total variation of gauge (6σ)	$= 6(\sigma^2_{repeat} + \sigma^2_{reprod})^{1/2}$
	= 6(0.0340)
	= 0.2040
Percentage over tolerance	= 25.5

In this case, gauge repeatability is very high compared to reproducibility, which indicates that the inspectors could not repeat measurements. A lack of training on how to use the measurement tool and the method of inspection including the basic initial requirements such as cleaning the surfaces, removing the burrs, and so on were identified as the major causes.

Variation in inspector capability in measuring the parts using the tool is the major cause for reproducibility variation. In the case of repeatability, it is mainly due to a lack of knowledge of measurement process. The general acceptable level of GRR values over tolerance is less than 10%.

Very similar results were observed for GRR studies conducted by the other participants. In all cases, gauge repeatability estimates were higher than the gauge reproducibility estimate. Management agreed to train all the inspectors and operators who are engaged in inspection within a time frame of 6 months.

Conclusion

The GRR estimates suggested by Montgomery and Runger (2) are biased. To make the estimates unbiased, d_2^* should be used instead of d_2. The confidence interval for GRR estimates based on the classical method can also be constructed. Care should be taken to avoid the mix-up of product variation with measurement error.

Acknowledgments

The author is extremely thankful to Michael E. Raynor, the referee, and Prof. V. Gopalan of Indian Statistical Institute for their valuable suggestions.

References

1. Montgomery, D. C., *Introduction to Statistical Quality Control*, 2nd ed., John Wiley and Sons, New York, 1991.
2. Montgomery, D. C. and Runger, G. C., Gauge Capability Analysis and Designed Experiments. Part I: Basic Methods, *Qual. Eng.*, 6, 115–135 (1993).
3. Rao, C. R., Mitra, S. K., Matthai, A., and Ramamurthy, K. G., *Formulae and Tables for Statistical Work*, Statistical Publishing Society, Calcutta, 1974.

About the Author: M. V. Asokan, masters degree holder in statistics, works presently as a Technical Officer in the SQC & OR Division of Indian Statistical Institute, Calcutta. He has organized several training programs in SQC/SPC, design of experiments, and related areas. He has published eight technical papers and articles.