Improved Gage R&R Measurement Studies

by **Donald S. Ermer**

Many manufacturers are using tools like statistical process control (SPC) and design of experiments (DoE) to monitor and improve product quality and process productivity. However, if the data collected are not accurate and precise, they do not represent the true characteristics of the part or product being measured, even if organizations are using the quality improvement tools correctly.

Therefore, it is very important to have a valid quality measurement study beforehand to ensure the part or product data collected are accurate and precise and the power of SPC and DoE are fully realized. Accuracy—in other words, no bias—is the function of calibration and is performed before a correct measurement study of the precisions of the gage and its operators.

In this two-part column, I will review the gage repeatability and reproducibility (R&R) study in the Automotive Industry Action Group (AIAG) manual[1] for its ability to determine the true capability of different parts of a measurement system. I'll use a geometrical approach to describe the components of the total measurement variance. This shows why the standard deviations or measurement errors of the equipment, appraiser and product in the AIAG method are not additive and cannot be compared directly in a ratio.

Part one of a two-part series on a geometrical approach.

In part two, which will appear in the May issue of *QP*, I will provide a worksheet for correctly executing a measurement process capability study that combines the advantages of my improved measurement study.

The Gage R&R Study In the AIAG Manual

The gage R&R study in the AIAG manual uses a data collection system that is well structured and very helpful in collecting the proper data. The data are then used to calculate the standard measurement errors, or standard deviations, of the equipment, appraiser and product. The total measurement error, or standard deviation, is then obtained by taking the square root of the sum of the squares of the equipment, appraiser and product standard deviations.

Next, the measurement ratios are calculated by comparing the equipment and appraiser standard deviations to the total measurement error or total standard deviation. These ratios are used to see how significant the different effects of the equipment, appraiser and product error variations are on the total measurement system.

Table 1 shows a summary of all the calculations used in this method, but only one standard deviation is used, instead of finding a 95% confidence interval, as in the AIAG manual.

Unfortunately, I find two errors in the AIAG R&R study. The first is a minor incorrect calculation of the part or product variation—in other words, there should be a correction factor that accounts for the variation induced by the measuring equipment. If this correction factor (although it may be very small) is not figured into the calculation, then equipment variation would be counted twice in the total variation.

The second and most significant error is the final variation ratios—percent equipment variation (%EV), percent appraiser variation (%AV) and percent part variation (%PV). These are calculated using standard deviations instead of variances (see the second column in Table 1). The results obtained exaggerate the proportional effects of the equipment, appraiser and part variation, as shown in the

TABLE 1 Summary of Equations (for One Standard Deviation) Used in the AIAG Gage R&R Study

Repeatability—equipment variation (EV)	$EV = \overline{\overline{WR}} * K_1$	$\overline{\overline{WR}}$ = Average of the within range of the trials of each part ($\overline{R}_e$)		
	%EV = 100 [EV/TV]	Trials (r) K_1	2 4.56	3 3.05
Reproducibility—appraiser variation (AV)	$AV = \sqrt{(\overline{X}_{Diff} * K_2)^2 - \frac{(E.V.)^2}{n * r}}$	$\overline{X}_{Diff}$ = Range of the operator averages (R_0)		
	%AV = 100[AV/TV] n = number of parts r = number of trials	Operators (k) K_2	2 3.65	3 2.70
Part variation (PV)	$PV = R_p * K_3$	R_p = Range of the part averages		
	%PV = 100[PV/TV]	Parts (n) K_3	5 2.08	10 1.62
Repeatability and reproducibility (R&R)	$R\&R = \sqrt{(EV)^2 + (AV)^2}$ %R&R = 100[R&R/TV]	%R&R<10%—process is capable 10%<R&R<20%—barely capable more than 20%—requires improvement		
Total variation (TV)	$TV = \sqrt{(R\&R)^2 + (PV)^2}$			

second column of Table 2.

Therefore, this incorrect type of study cannot provide an index of whether the components of the measurement process are capable for the part or product under study.

The Correct Calculation

The measurement equipment error and part variation can be related by the following equation, assuming measurement error is independent of the part variation within a range defined by the natural process limits for the specified product:

$$\sigma_m^2 = \sigma_p^2 + \sigma_e^2$$

where:

σ_m^2 = variance of actual product measurement [%PV],

σ_p^2 = true variance of product and

σ_e^2 = variance of measurement equipment error [%EV]

This relationship can also be represented by a right triangle, as shown in Figure 1. For example, if $\sigma_m = 5$, $\sigma_e = 3$ and $\sigma_p = 4$, then $3^2 + 4^2 = 5^2$, and $\frac{3^2}{5^2} = \frac{9}{25} = 0.36$, which is not equal to $\frac{3}{5} = 0.60$; and $\frac{4^2}{5^2} = \frac{16}{25} = 0.64$, which is not equal to $\frac{4}{5} = 0.80$.

This is a simple illustration of some of the misleading results for the final variation ratios in the AIAG method. Thus, a unit change in either the true product standard deviation (σ_p) or the standard deviation for measurement equipment error (σ_e) will not result in a unit change in the standard deviation for the actual product measurements (σ_m). On the other hand, one unit change in the true product variance (σ_p^2) or measurement equipment error variance (σ_e^2) will respond to one unit change in the variance of actual product measurement (σ_m^2), because $\sigma_m^2 = \sigma_p^2 + \sigma_e^2$, or $\sigma_m \neq \sigma_p + \sigma_e$.

This shows part of the results of the gage R&R report in the AIAG method is incorrect. Therefore, a person doing the study should use variances to find the true individual percentage R&R ratio values instead of the standard deviations, where:

$$\hat{\sigma}_e = \frac{\overline{WR}}{d_{2,e}}$$

$$\hat{\sigma}_m = \frac{R_p}{d_{2,m}^*}$$

and

$\overline{WR}$ = average of all the within ranges of the trials for each part,

R_p = range of the actual parts averages,

$d_{2,e}$ = 1.128 and 1.6926 for the second and third trials, respectively, from Table 3 (assuming n = 10 parts and k = 3 appraisers),

$d_{2,m}^*$ = 2.48 (see top of Table 3 for k=1 and 5 parts) and 3.18 (see bottom of Table 3 for k=1 and 10 parts); see complete Table 3 for other d_2* values.[2, 3]

It should be noted that d_2* values should be used when k = 1 (see first row of Table 3), because of the significant difference from assuming k>25.

TABLE 2 Summary of Results From Different R&R Studies For the Same Measurement Data

Measurement unit analysis I	AIAG method (incorrect percentages) II	Analisis of variance (uses all the data) III	New method (using d_2* and variances) IV
EV (σ_e)	3.783	4.41	3.710
AV (σ_0)	4.286	3.50	4.293
R&R = (EV+AV)[1/2]	5.717	5.63	5.674
PV (σ_m)	23.45	23.05	23.43
TV	24.13	23.71	24.11
%(EV/TV)	15.68%	**3.45%**	**3.09%**
%(AV/TV)	17.76%	**2.18%**	**2.45%**
%(PV/TV)	97.18%	**94.37%**	**94.46%**
TOTAL	"130.62%"	**100.00%**	**100.00%**

EV = equipment variation AV = appraiser variation R&R = repeatability and reproducibility

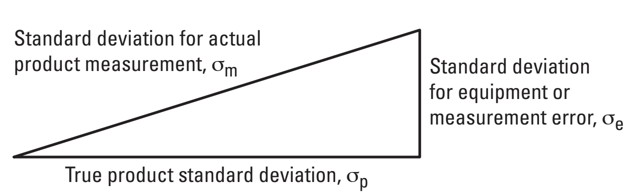
PV = part variation TV = total variation

The Results of Different R&R Studies

Table 2 shows the results of using different gage R&R studies for the same set of measurement data, where the analysis of variance (ANOVA) method is the most accurate, since it uses all the data, not just the ranges. The last column (IV) shows the correct percentage variations calculated from the new method I've partly described. It should be noted that with the new method, the repeatability, reproducibility and part variation percentages add up to 100%, which is not true with the AIAG method.

Comparing the AIAG R&R method with the ANOVA method (for the random effects model) and the new method, we can see the AIAG method exaggerates the effect of the %EV,

FIGURE 1 The Relationships Among the Standard Deviations



%AV and %PV. That is, the %EV, %AV and %PV as obtained in the AIAG R&R study are incorrectly greater than the actual results, as shown in the second column of Table 2. This is due to the incorrect approach of comparing the standard deviations to the total standard deviation vs. the correct approach of using the additive law of variances.

This mistake in the AIAG R&R study could lead to a conclusion that the parts of the measurement process are incapable when they are actually capable.[4] Also, mistakes such as these may mislead an organization to try to improve a capable measurement process, when it should actually reduce the process or product variance. The results of the ANOVA method and new method, however, agree closely. This can be seen in columns three and four of Table 2.

Change Is Needed

AIAG R&R methods may be misleading and should be modified. Reliable measurement data and analysis of those data are important. Part one of this column will help conscientious organizations further improve the quality of their products and the productivity of their processes.

In part two, I will introduce the appraiser variation (AV) as the third component of the total product measurement variation and will give an example of a complete and correct gage R&R measurement study.

REFERENCES

1. L.A. Brown, B.R. Daugherty and V.W. Lowe, *Measurement Systems Analysis*, third edition, Auto Industry Action Group, 2003.
2. Acheson J. Duncan, *Quality Control and Industrial Statistics*, fifth edition, Richard D. Irwin Inc., 1986.
3. Donald S. Ermer and P.E. Prond, "A Geometrical Analysis of Measurement System Variations," *ASQC Annual Quality Congress Transactions*, 1993.
4. Donald S. Ermer, "Pythagorean Theorem to the Rescue or Reliable Data Is an Important Commodity," *The Standard*, ASQ Quality Measurement Division, Winter 2000/Spring 2001.

BIBLIOGRAPHY

Jay Bucher, *The Metrology Handbook*, ASQ Quality Press, 2004.

Ronald D. Snee, "Are You Making Decisions in a Fog?" *Quality Progress*, December 2005.

Philip Stein, "All You Ever Wanted to Know About Resolution," *Quality Progress*, July 2001.

TABLE 3 Constants for Converting a Range To a Standard Deviation

Values of d_2^*, d_2, and d_3						
	Sample size (n)					
Number of subgroups (k)	2	3	4	5	6	7
1	1.41	**1.91**	2.24	**2.48**	2.67	2.83
2	1.28	1.81	2.15	2.40	2.60	2.77
3	1.23	1.77	2.12	2.38	2.58	2.75
4	1.21	1.75	2.11	2.37	2.57	2.74
5	1.19	1.74	2.10	2.36	2.56	2.73
6	1.18	1.73	2.09	2.35	2.56	2.73
7	1.17	1.73	2.09	2.35	2.55	2.72
8	1.17	1.72	2.08	2.35	2.55	2.72
9	1.16	1.72	2.08	2.34	2.55	2.72
10	1.16	1.72	2.08	2.34	2.55	2.72
11	1.16	1.71	2.08	2.34	2.55	2.72
12	1.15	1.71	2.07	2.34	2.55	2.72
13	1.15	1.71	2.07	2.34	2.55	2.71
14	1.15	1.71	2.07	2.34	2.54	2.71
15	**1.15**	1.71	2.07	2.34	2.54	2.71
16	1.15	1.71	2.07	2.34	2.54	2.71
17	1.15	1.71	2.07	2.34	2.54	2.71
18	1.15	1.71	2.07	2.33	2.54	2.71
19	1.14	1.70	2.07	2.33	2.54	2.71
20	1.14	1.70	2.07	2.33	2.54	2.71
25	1.14	1.70	2.07			
30	1.14	1.70				
50	1.13					
d_2	1.128	1.6926	2.0588	2.3258	2.5344	2.7044
d_3	0.8525	0.8884	0.8798	0.8641	0.8480	0.8332

Source: A.J. Duncan, "The Use of Ranges in Comparing Variabilities," *Industrial Quality Control*, Vol. 40, No. 5, February 1955; No. 8, April 1955.

Note for k = 1 and n = 10, d_2^* = 3.18

DONALD S. ERMER *is the Procter & Gamble professor emeritus of total quality at the University of Wisconsin-Madison, where he also earned his doctorate in mechanical engineering. Ermer is an ASQ Fellow, is chair of Madison Section 1217 and received ASQ's 1997 Edward Oakley Award and 2000 National Grant Medal.*

Please comment

If you would like to comment on this article, please post your remarks on the *Quality Progress* Discussion Board at www.asq.org, or e-mail them to editor@asq.org.