What Is the Difference between ARM and ANOVA?

Stephan Conrad, TEQ® Training & Consulting GmbH

Since the publication of the 4th edition of the AIAG Core Tool MSA reference manual, there is one modification that has given the participants of our seminars plenty to talk about. This manual recommends the ANOVA method for the calculation of Gage R&R results. It considers the average range method to be unsuitable which causes huge problems for users of "download-Excel-tables-from-the-Internet" tools. Whereas ARM seemed so simple to interpret and was even easily comprehensible with the help of a calculator, ANOVA hides behind a whole lot of abstract, complex formulas.

However, in the meantime, almost every OEM demands the application of analyses of variances. Even larger suppliers such as Bosch completely adapt their new Bosch booklet 10 to ANOVA. This booklet actually refers to ARM as "outdated". It seems like there is no way around it – you have to deal with this calculation method.

An Illustrative Explanation, Just for Once …

Well, not everybody wants to pore over formulas and work through incredibly long multiple sums. But how can we explain the approach and particularly the advantages of ANOVA? Let's have a try. And to put all critics on the alert by now – the author does not attach any importance to mathematical precision but tries to explain these phenomena comprehensibly and wants to underline the special features of these methods. As an example, the graphics used in the following will hardly be able to illustrate that you have to add variances instead of standard deviations. This is why I want to put something straight – never perform an addition, subtraction or whatsoever type of operation with standard deviations. Not for nothing refers ANOVA to "analysis of variance". So the calculation is always based on the squares of standard deviations that are also called variances!

First of all ARM again

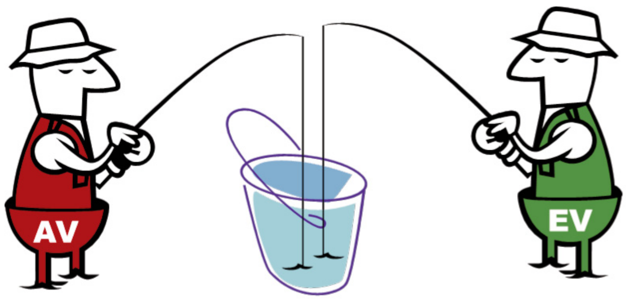
As described in the "ARM, ANOVA and All the Rest…" article published in PIQ 2/2011, when using the average range method, you just read repeatability and reproducibility directly from a very well-arranged table of measured values.

Appraisers repeat their own measurement and calculate ranges. There are three appraisers and each appraiser measures 10 parts twice. So you obtain 10 ranges per operator. Now you calculate the average range for each appraiser trial. You obtain three average ranges, one for each appraiser. Use these three ranges to calculate the mean value of all ranges. Unfortunately, the result is a range, of course, and not a standard deviation. For this reason, you have to add correction factor K1. The calculation result specifies the repeatability or equipment variation (EV).

Now, each of these three appraisers has taken 20 measurements whose measured values should actually be equal on average. In order to find differences, you just take the averages of the three operators calculated from the respective 20 measurements and subtract the smallest average from the largest average. The result is a range one again, so you have to apply correction factor K2 leading to the reproducibility or appraiser variation (AV).

The main problem with this approach is the fact that these two variation components are calculated based on two clearly outlined phenomena. Everything that does not fit this pattern will not be considered and detected. Moreover, including correction factors in the calculation is actually a popular "source of errors" – whether you are aware of it or not.

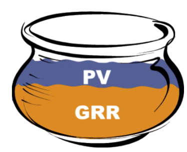
I have got a picture in my head that illustrates the situation quite well. The average range method is like two anglers that both have their own special baits in order to catch nothing but the fish they want to have.



And now ANOVA …

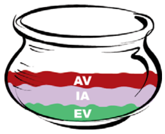
So this is the reason why we use the calculations according to ANOVA. From now on, I will merely focus on a "good illustration" of this approach. First, we evaluate the total variation of type-2 study. We thus take all 60 measured values, lump them together in a pot and obtain the "total variation" of this study type. MSA 4 also refers to this variation as "total variation TV".

Roughly speaking, part of the variation comes from the measurement system whereas the other part is caused by the manufacturing process from which we actually borrowed our 10 parts. The MSA calls these components GRR and "part variation PV".



Basically, we are only interested in GRR, so this is why we take out the part variation PV and put it aside. Maybe we might still need this soup of part variation later on.

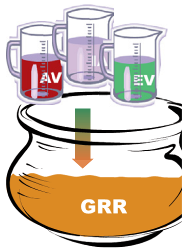
Now we have actually finished because all that remains in the pot is the GRR for which we were actually looking. However, the truth is that this issue now starts getting exciting. We can analyze the measurement system variation and find out the components of which it consists.



First we calculate the reproducibility AV from the averages of the appraiser trials. The approach is similar to the average range method but it immediately leads to the correct result and does not require any K-factor. In the next step, we scoop the reproducibility out of the pot – it gets emptier.

In case the averages of the appraiser trials vary around AV, the variation of the three appraiser averages of each part must not be any wider. Otherwise, the additional variation is caused by the interaction IA between appraiser and part. If we are able to detect this interaction, we scoop it out of the pot, too. This is another characteristic feature of the ANOVA method; ARM is not even able to consider the interaction component!

Now, the pot is even emptier and only a single variation component remains. So this is where we have come across the biggest and most important difference. All the rest is referred to as the repeatability or equipment variation EV, i.e. ANOVA empties the pot to the last drop and there is nothing that is left over, nothing it forgets to consider. In comparison to the "fishing for variation" method according to ARM, we cannot overlook anything; ANOVA takes any available variation component into account and evaluates it.



However, the conclusion is not that ANOVA always provides worse results than the average range method. In most cases, the plain average range method quite simply leads to totally different and sometimes even considerably larger results than ANOVA.

Now that we know the ingredients of our soup, we pour all the three components back into the pot and refer to this soup as GRR.

$$SS_P = \sum_{i=1}^{n} \frac{x_{i\bullet\bullet}^2}{kr} - \frac{x_{\bullet\bullet\bullet}^2}{nkr}$$

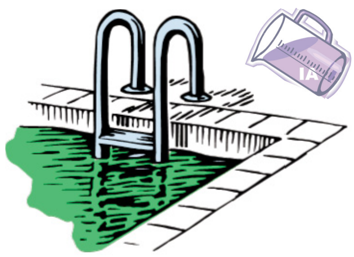
$$SS_A = \sum_{j=1}^{k} \frac{x_{\bullet j\bullet}^2}{nr} - \frac{x_{\bullet\bullet\bullet}^2}{nkr}$$

$$SS_{AP} = \sum_{i=1}^{n} \sum_{j=1}^{k} \frac{x_{ij\bullet}^2}{r} - \frac{x_{\bullet\bullet\bullet}^2}{nkr} - SS_P - SS_A = \sum_{i=1}^{n} \sum_{j=1}^{k} \frac{x_{ij\bullet}^2}{r} - \sum_{i=1}^{n} \frac{x_{i\bullet\bullet}^2}{kr} - \sum_{j=1}^{k} \frac{x_{\bullet j\bullet}^2}{nr} + \frac{x_{\bullet\bullet\bullet}^2}{nkr}$$

$$SS_E = \sum_{i=1}^{n} \sum_{j=1}^{k} \sum_{m=1}^{r} x_{ijm}^2 - \frac{x_{\bullet\bullet\bullet}^2}{nkr} - SS_P - SS_A - SS_{AP} = \sum_{i=1}^{n} \sum_{j=1}^{k} \sum_{m=1}^{r} x_{ijm}^2 - \sum_{i=1}^{n} \sum_{j=1}^{k} \frac{x_{ij\bullet}^2}{r}$$

If you want to learn more about this approach in detail, you are welcome to look at the calculation in the AIAG Core Tool MSA reference manual. Chapter III section B shows a calculation example and appendix A explains the theory on which the calculation is based. This topic is also perfectly presented in the appendix of Bosch booklet 10 and last but not least in the "Measurement Process Qualification" reference book by Edgar Dietrich and Alfred Schulze published by Carl Hanser Verlag.

Incidentally, now it becomes clear what the "pooled" interaction is all about. If you find out that the IA measuring cup is almost empty compared to the EV pot - so that it is not really worth mentioning an interaction - you can just pour the interaction back "into the pool" and spoon the vanishingly little proportion of interaction out together with the repeatability.



In case you found this article comprehensible, do not tell your colleagues. They are welcome to keep struggling with these formulas. However, we are interested in your opinion. Discuss this topic with us! Send an e-mail to stephan.conrad@teq.de to contact the author.