Why Use Ranges?

What they didn't teach in your stat class

Donald J. Wheeler

Last month in "The Analysis of Experimental Data" I presented a method for analyzing experimental data that was built on the use of the range statistic as a measure of dispersion. In this day of computers and software, why should we even consider using ranges in our analysis of experimental data? Wouldn't other, more efficient measures of dispersion do a better job? Since these questions can be a barrier to the effective analysis of data, they deserve to be answered.

Last month I showed how to use the Analysis of Means (ANOM). When Ellis Ott introduced this technique in 1967 he used the average range as the basis for obtaining the ANOM detection limits. When Ed Schilling extended Ott's technique in 1973 he also used the average range as the basis for obtaining the ANOM detection limits. In these early papers the scaling factors given were merely upper bounds on the exact values. In 1974 Lloyd Nelson recomputed these upper bounds in order to obtain more precise ANOM detection limits. However, in his work Lloyd shifted from using the average range to using a different dispersion statistic. Lloyd used the Root Mean Square Within. We will return to this in a moment. In 1975 Ellis Ott used the average of upper and lower bounds on the scaling factors to obtain sharper more precise ANOM detection limits based on the average range statistic. Finally, in 1982 Peter Nelson computed exact critical values for ANOM based on the Root Mean Square Within statistic. At the suggestion of this author Peter nelson extended these tables in 1993. Thus, from the beginning, much effort was spent in refining and obtaining exact scaling factors for ANOM. Along the way a shift occurred in the dispersion statistic used in the computation. Since this shift is not a trivial shift, and since it has some unintended consequences, we need to compare the use of these two measures of dispersion.

THE ROOT MEAN SQUARE WITHIN STATISTIC

For over 80 years the Mean Square Within has been the standard yardstick for characterizing the within-subgroup dispersion in the Analysis of Variance (ANOVA). Given k subgroups of size n, the Mean Square Within is simply the average of the k subgroup variance statistics:

$$\text{Mean Square Within} \quad = \quad MSW \quad = \quad \overline{s^2}$$

The reason that the Mean Square Within works so well with ANOVA is due to the nature of the ANOVA technique. In ANOVA we essentially compute a signal-to-noise ratio. While the Mean Square Within provides the most efficient unbiased estimate of the noise component, the signal component is found by *squaring* all the potential signals and adding them up. ANOVA then compares the average of the squared signals with the Mean Square Within. This computation of a signal-to-noise ratio has been found to be very robust and satisfactory in over 80 years of use.

Because of this success, it is almost natural for statisticians, who are extensively trained in ANOVA techniques, to immediately think of using the *Root Mean Square Within* for applications requiring an estimate of the standard deviation parameter.

$$\text{Root Mean Square Within} \ = \ \sqrt{MSW} \ = \ \sqrt{s^2}$$

While the Mean Square Within is completely appropriate for ANOVA, the Root Mean Square Within has a drawback that makes it inappropriate for ANOM. While the Mean Square Within is robust when it is used with ANOVA, the Root Mean Square Within suffers from a lack of robustness.

ROBUSTNESS

The comparison made in ANOM is different from the comparison made in ANOVA. In ANOM the potential signals are simply plotted one by one, and those that cross over the detection limit are those that are likely to be real. Thus, in ANOM we filter out the probable noise by the computation of the detection limit. And as you can see from the brief history given above, much effort has been put into the problem of sharpening up the computation of these detection limits.

So, why not use the Root Mean Square Within to compute ANOM limits? Because the Root Mean Square Within is not robust. When the subgroups do not all display the same amount of variation the Root Mean Square Within will be more inflated than will the Average Range. This will result in limits that are wider than they should be, which will make the analysis less sensitive than intended.

To illustrate this problem I will use an example from one of my clients. These data consist of ten subgroups of size two. The Averages, Ranges, and Variances for each subgroup are listed in Figure 1.

Subgroup	1	2	3	4	5	6	7	8	9	10
Values	20	20	25	45	50	20	15	15	35	45
	15	25	25	40	50	20	10	10	40	20
Averages	17.5	22.5	25.0	42.5	50.0	20.0	12.5	12.5	37.5	32.5
Ranges	5	5	0	5	0	0	5	5	5	25
Variances	12.5	12.5	0	12.5	0	0	12.5	12.5	12.5	312.5

Figure 1: The Gauge Study Data

Even a cursory glance will show that the last subgroup has a different amount of variation from the others. To see the impact of this subgroup we shall compute measures of dispersion using the first nine subgroups and then using all ten subgroups.

The first nine subgroups have an Average Range of 3.333. Yet the Average Range for all ten subgroups is 5.5. Thus, the excessive variation in the last subgroup inflates the Average Range by 65%.

The first nine subgroups have an Root Mean Square Within of 2.89. Yet the Root Mean Square Within for all ten subgroups is 6.22. Thus, the last subgroup inflated the Root Mean Square Within by 115%, which is almost twice the inflation shown by the Average Range. This tendency toward excess inflation is inherent in the Root Mean Square Within. By squaring the

standard deviations, averaging them, and then finding the square root, the Root Mean Square Within automatically gives more weight to the larger values.

Thus, the Root Mean Square Within is not robust to differences in variation. This lack of robustness is not so much of a problem with ANOVA as it is with ANOM. This is because in ANOVA we are squaring the signals as well as squaring the noise. (When the signals are larger than the noise, their squares increase even more rapidly than the contamination in the Mean Square Within.) However, when we are computing limits to use with the raw signals themselves this lack of robustness can result in detection limits that are more inflated than those found using the Average Range. It is this lack of robustness that makes the Root Mean Square Within an inappropriate measure for use with ANOM or process behavior charts.

This subtle point was missed by those that shifted to using the Root Mean Square Within with ANOM. While they were busy refining the upper bounds provided by Ott in order to sharpen up the detection limits, their shift to using the Root Mean Square Within ignored its lack of robustness. The use of the Average Range provides additional robustness to the analysis that is missing when the Root Mean Square Within is used. Figure 2 shows the ANOM plots for Figure 1 with detection limits based on each measure of dispersion.
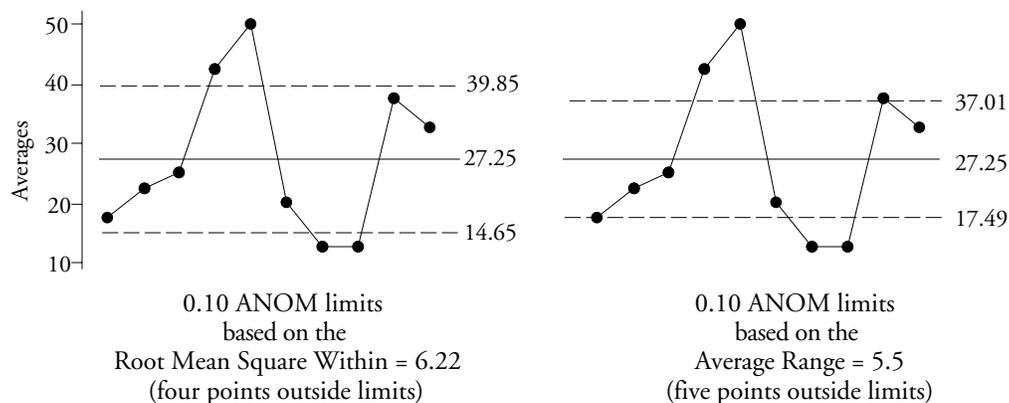

0.10 ANOM limits
based on the
Root Mean Square Within = 6.22
(four points outside limits)

0.10 ANOM limits
based on the
Average Range = 5.5
(five points outside limits)

Figure 2: The Average Range Is More Robust than the Root Mean Square Within

So, you have paid good money to conduct the experiment and collect your data, does it make sense to use an analysis that will miss some of the signals you have paid to discover?

EQUIVALENCE

"Okay, the Root Mean Square Within is not robust for ANOM. But I was taught that the range is inefficient, so what about using the standard deviations instead of the variances?"

While it is true that the range is an inefficient summary for a large data set, this does not present a problem for ANOM and process behavior charts because they are built on using small sets of data. However, let us consider what would happen if we were to use standard deviations. If we return to the data of Figure 1 and consider the subgroup standard deviations we have the values shown in Figure 3.

The first nine subgroups have an Average Standard Deviation of 2.36. Yet the Average Standard Deviation for all ten subgroups is 3.89. Thus, the excessive variation in the last

subgroup inflates the Average Standard Deviation by 65%. This is exactly the same as the inflation seen with the Average Range described in the previous section.

Subgroup	1	2	3	4	5	6	7	8	9	10
Values	20	20	25	45	50	20	15	15	35	45
	15	25	25	40	50	20	10	10	40	20
Std. Dev.	3.54	3.54	0	3.54	0	0	3.54	3.54	3.54	17.68

Figure 3: Standard Deviations for the Gauge Study Data

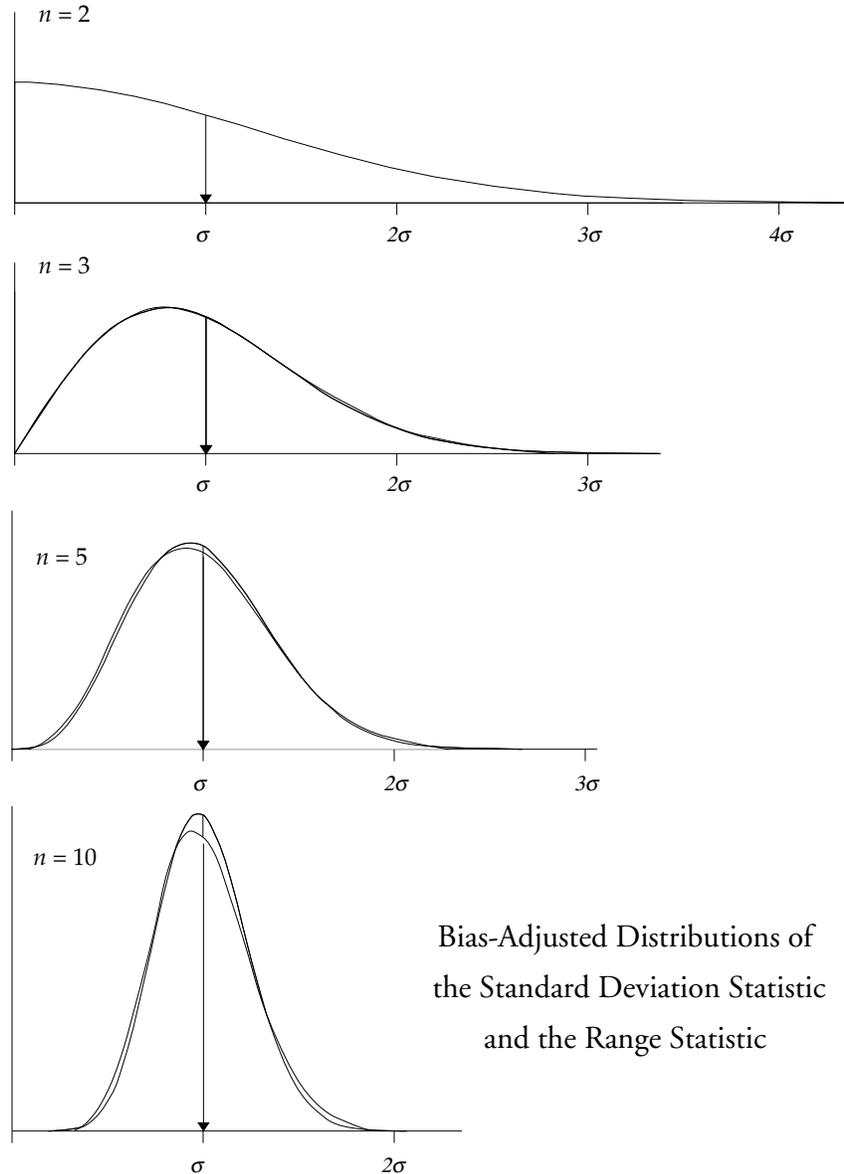Bias-Adjusted Distributions of the Standard Deviation Statistic and the Range Statistic

Figure 4: The Range and Standard Deviation are Equivalent for Small n

As I showed in my book *Advanced Topics in Statistical Process Control*, the range is essentially equivalent to the standard deviation statistic when the subgroup size is smaller than 15. The

graphs shown in Figure 4 superimpose the distributions of both statistics after adjusting for their different biases. In the case of $n = 2$ there is only one curve because of the deterministic relationship between the range and the standard deviation. For the larger values of n there are two curves shown. After adjusting each distribution for the average bias of the statistic we end up with curves that are so close together that it is an eye test to separate them. In each case the differences between the two distributions shown for each value of n in Figure 4 are too small to be of any practical interest.

Not only are the distributions essentially the same, but the individual statistics for a given subgroup are highly correlated with each other when n is less than 15. Figure 5 shows the correlations between the bias-adjusted ranges (on the x-axis) and the bias-adjusted standard deviation statistics (on the y-axis) for each of four sets of 100 subgroups.
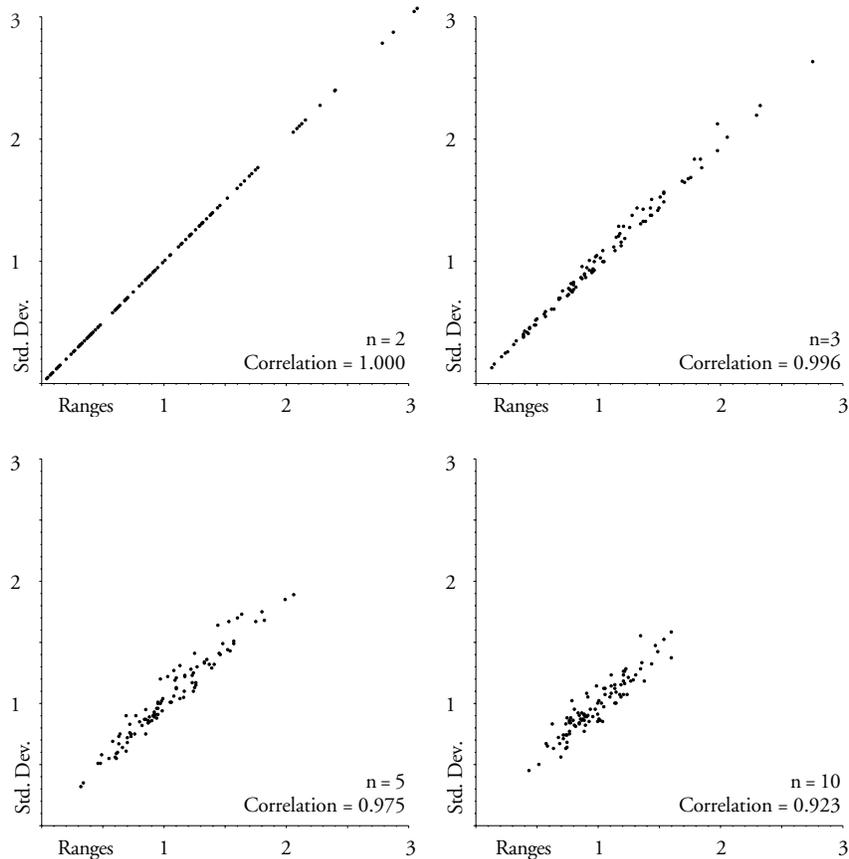

Figure 5: Correlations Between Ranges and Standard Deviations

The correlation is 1.00 when $n = 2$ and it is 0.996 when $n = 3$. Then, as n increases from 3 to 15 the correlation between the range and the standard deviation drops approximately 0.01 with each unit change in n. Thus, as long as your subgroup size is less than 12, the range and the standard deviation will have a correlation greater than 90%.

The fact that these statistics are equivalent for small subgroup sizes means that the choice between using a range and using a standard deviation statistic to measure the variation within the subgroups is merely a personal preference rather than being a matter of mathematical

efficiency. Some people actually prefer complexity. Regardless of your personal preference, both the Average Standard Deviation and the Average Range will work equally well with small subgroups. However, the Average Range has one advantage that the Average Standard Deviation lacks.

TRANSPARENCY

On a completely different level there is another reason to use the range. For over 40 years I have had the job of teaching statistical techniques to non-mathematicians. As part of this job I have had the opportunity to look in into student's faces as explanations are given and I have discovered that even among those who know how to compute a standard deviation statistic there is less resistance when I use a range. A range is an intuitive and transparent measure of dispersion. The standard deviation statistic is neither.

For example, I find that people seem to understand the average as the center of mass for the data. I suspect that this is because balance is one of those things we learn as a toddler. Yet when I extend this physical analogy and explain the variance statistic as the rotational inertia for the data I tend to get blank looks. I suspect that this is because we are not ordinarily concerned with rotational inertia except on those occasions when we slip on the ice and lose our balance.

So, whenever I use the range in my analysis I find that the explanation is generally met with understanding and acceptance. On the other hand, when I use the standard deviation in my analysis, I find that the explanation gets derailed. As soon as I mention the standard deviation half the class gets a blank look in their eyes and the other half starts trying to figure out how the standard deviation fits into the analysis. Either way, the explanation has been derailed and turns into gobbledygook. Since the most important part of any analysis is the communication of the results, it is important to avoid gobbledygook, and one of the easiest ways to do this is to avoid using the standard deviation statistic. After all, even if someone does understand the concept of rotational inertia, they are still unlikely to understand exactly what the square root of rotational inertia represents.

SUMMARY

Thus, there are three reasons I use the Average Range with ANOM and process behavior charts rather than the Root Mean Square Within. The first of these is mathematical robustness; the second is the practical equivalence of the range and the standard deviation statistic for small subgroup sizes; and the third is the intuitive and transparent nature of the range. In keeping with the principle that the best analysis is the simplest analysis that allows you to discover the signals contained within your data, the range-based Analysis of Means is not only robust, intuitive, transparent, and easy, but it is also hard to beat.

(This article is excerpted from Dr. Wheeler's latest book, *Analyzing Experimental Data*, © 2013 SPC Press and is used with permission.)