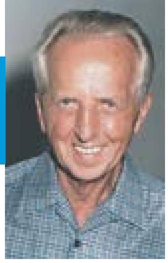

ACADEMIC STUDY
Dr. David Kirk | Coventry University

Shot Peening Statistics

INTRODUCTION

The purpose of this article is to assist readers in understanding the increasing number of applications of statistics in shot peening. Mathematics here is kept as simple as possible. The worst abuse of statistics occurs when measurements are simply entered into a formula which is not appropriate. Statistical programs are now routinely available, e.g., within Excel.

Many shot peening factors vary including shot particle diameter, air pressure, wheel speed and Almen gauge parameters. Best practice demands that measurement values, i.e., data, are carefully stored and accessible. Every piece of data can be regarded as being a result from an experiment and has lasting value. It is regrettable that some companies discard data after it has served its immediate purpose.

Statistics is the science of making decisions in the face of uncertainty. We cannot know, for example, what exactly the arc height of an individual peened Almen strip will turn out to be. This is in spite of our best efforts. Random variation of measurement factors will always occur and there may also be systematic variation—as, for example, when supplied air pressure falls steadily.

METHODS OF ANALYSING DATA

The most commonly used methods of analysing data are either pictorial or arithmetical.

Pictorial Methods

Bar charts and histograms are familiar ways of displaying collections of data values. Playfair introduced bar charts in 1781 and histograms were introduced by Pearson in 1891. Table 1 is a hypothetical data set for thickness measurements on Almen strips.

Table 1. Hypothetical set of thickness data values for a box of Almen A strips.

Thickness band - mm	Number of strips
A 1.27-1.275	6
B 1.28-1.285	15
C 1.29-1.295	40
D 1.30-1.305	30
D 1.30-1.305	9
Total	100

Using the pictorial Bar chart method with Table 1 data we get fig.1.

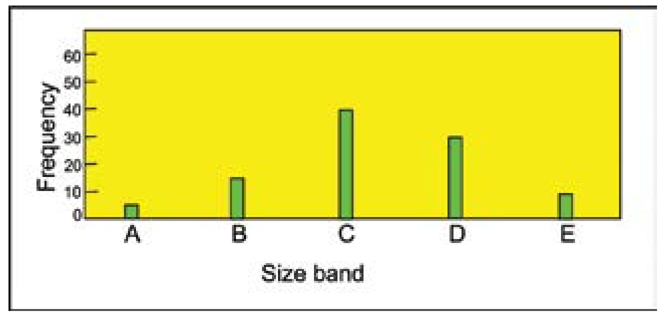

Fig.1. Bar Chart of Table 1 data.

Using the pictorial histogram method with Table 1 data we get fig.2.

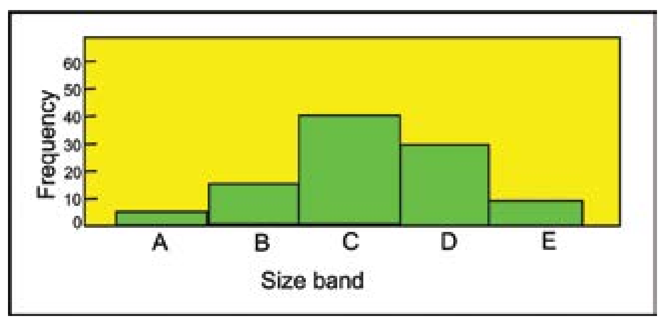

Fig.2. Histogram of Table 1 data.

A comparison of the same data, presented in figs.1 and 2, reveals the advantages of histograms. The principal advantage is that the size band width indicates the variation within each band. It is perhaps surprising that it took over a century for histograms to largely supersede bar charts.

Arithmetical Location Methods

Arithmetical methods produce quantities that summarize the data. Each quantity is then properly called a "statistic".

The **mean** is by far the most important commonly used measure of location. To obtain the mean we simply add up all the values in the data set and divide by the number of values in the data set. The term "average" is synonymous with "mean".

The **median** is the magnitude for which half of the data values are less than the median and half are greater than the median. It is meaningful if the frequency plot is severely skewed.

The **mode** is the value of the variable that occurs with the greatest frequency. The midpoint of the tallest box gives a good estimate of the mode. For the data given in table 1 this is 1.2925 mm (the middle of size band C in Table 1).

When the size distribution of data values is roughly symmetric, the mean, median and mode values will be very close together. If, however, the distribution is very skewed they will have quite different values. Fig.3 is an example of a severely skewed distribution.

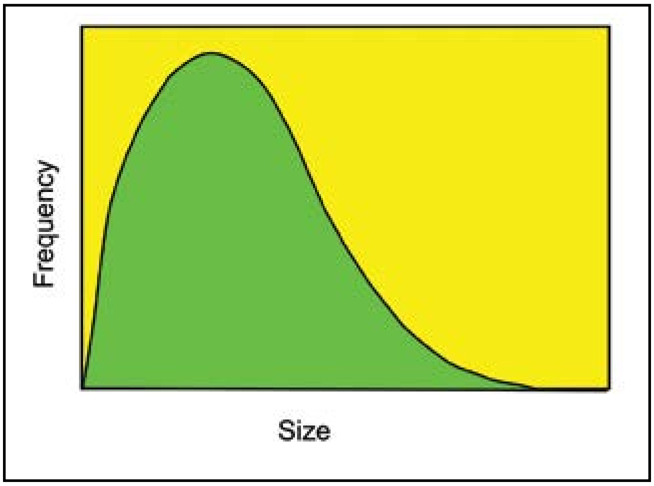

Fig.3. Skewed size frequency curve.

Arithmetical Variability Methods

It is often important to be able to quantify the variability of the data within a set. The simplest method is the **range**; this is the difference between the largest and smallest values in the data set. However, there are strong practical reasons for preferring a statistic called the "**variance**" , or its square root, which is called the "**standard deviation**". The mathematical bases for variance and standard deviation are of very limited interest to most shot peeners. Consider, however, a different situation. Imagine that we are trying to determine whether or not a set of newly minted coins are biased. Using the "heads or tails" approach, tossing a single coin would not allow any conclusion to be drawn. If two coins were tossed there are three possible outcomes—two heads, two tails or one head and one tail. The outcome would give a faint indication of coin bias. Tossing three coins would give a much better indication. A four-heads outcome would arouse significant doubt as to lack of coin bias. The moral is that the larger the number in any data set the lower will be its variance. An example of applying arithmetical variability methods is, however, given as follows:

Find the range, variance and standard deviation of these six measurements.

0.9, 1.3, 1.4, 1.2, 0.8 and 1.0.

Note that both variance and standard deviation values are easily calculated using readily available programs. For example, using Excel. Enter the six values of this data set into A1 to A6. Then highlight any other box. In the formula bar type = STDEV.P(A1:A6) and press Enter. The standard deviation value then shows up immediately as 0.216.

Excel results for this data set:
Range = 1.4 – 0.8 = **0.6**
Variance, s^2 = 0.0467
Standard Deviation, s = 0.216

ACCURACY AND PRECISION

Having been able to assess data set location and its variability, attention can now be turned to its accuracy and precision. Figs.4 to 7 illustrate the significance of the parameters of these normally distributed Almen arc heights. Fig.4 shows the ideal situation where (a) the average of the measurements coincides with the true arc height and (b) the measurements have a low variability, ranging from a to b.

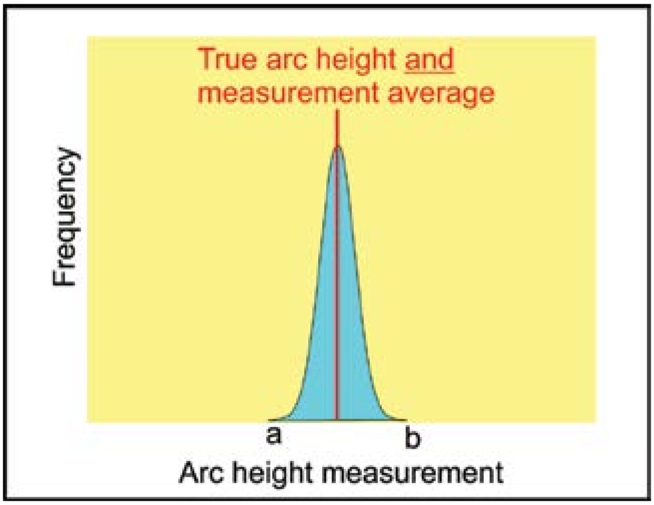

Fig.4. Good accuracy and good precision.

For fig.5 (page 30), the average of the measurements is substantially different from the true arc height—indicating poor accuracy. Bias is the name given to the difference between any true value and a measurement mean. The variability could, however, have been good—as good as that shown in fig.4—indicating good precision.

For the situation shown in fig.6, the accuracy is good since the measurement average is the same as the true value. The measurements do have considerable variability thus indicating poor measurement precision.

The worst case scenario is indicated in fig.7 where both accuracy and precision are poor.

COMPARISON OF DATA SETS

Table 2 illustrates how comparison statistics can be employed. For this example, two sets of Almen strips, A and B, from the same box, were peened. Each strip was given the same

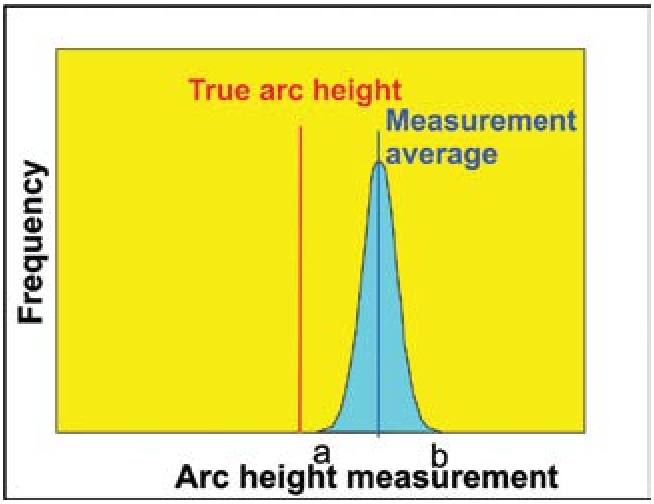

Fig.5. Poor accuracy but good precision.

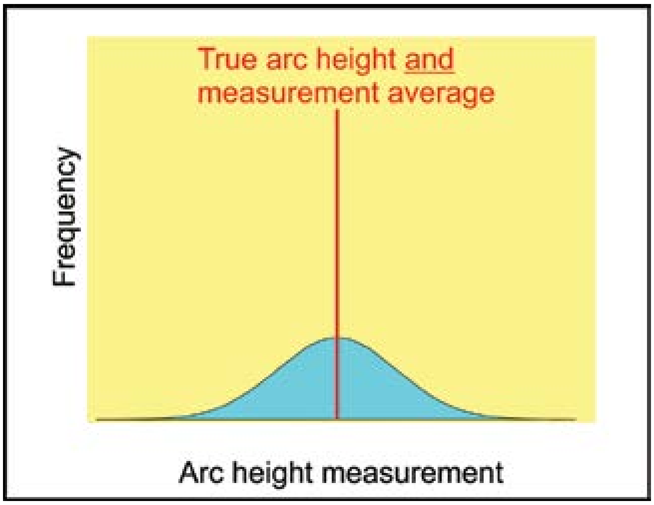

Fig.6. Good accuracy but poor precision.

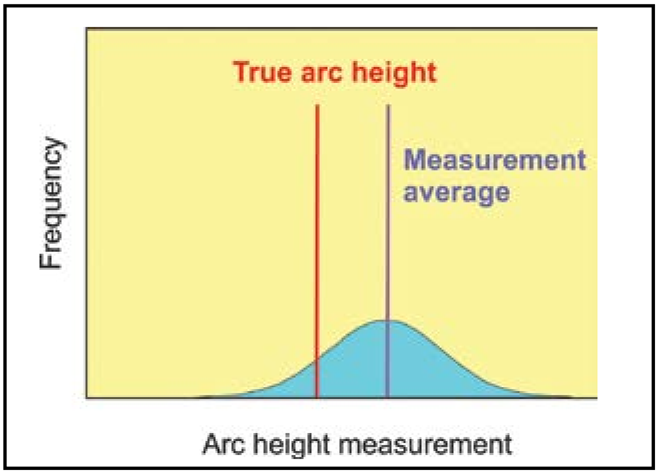

Fig.7. Poor accuracy and poor precision.

nominally identical exposure and intensity. Measured arc heights varied, with those in Set A being much less variable than those in Set B. The reasons will be discussed later in the article.

Table 2. Variability comparison for two sets of peened Almen strips.

Strip No.	Arc heights (inch x 1000)	
	Set A	Set B
1	6.2	6.3
2	6.3	6.5
3	6.3	5.9
4	6.2	6.7
5	6.5	6.0
6	6.3	5.9
7	6.3	6.4
8	6.4	6.3
9	6.2	6.2
10	6.3	6.5
11	6.3	6.7
12	6.1	5.9
Mean	**6.30**	**6.30**
Standard deviations	**0.1**	**0.30**

Table 3 presents a useful quantification of relative variability for the two sets of strips.

The magnitude of the standard deviation allows us to predict the probability of a future single measurement being away from the mean. This probability is stated in Table 3.

Table 3. Probability of a new measurement's value relative to the mean.

Number of standard deviations away from the mean	Probability of obtaining a new measurement value
1	**One in three**
2	**One in twenty**
3	**One in four hundred**

The universally accepted values given in Table 4 can be applied to the measurement values given in Table 2. Remember that "probability" is not the same as "certainty". For one standard deviation away from the mean, Set A contains four measurements—1,5,8 and 12—which just happens to be "one in three". For Set B there are five values—3,4,6,11 and 12—which is less close to "one in three". For two standard deviations away from the mean Set A has just strip 5 outside—close to the probability of "one in twenty". Set B doesn't have any—still not too far from the probability of "one in twenty". For three standard deviations from the mean neither set has a strip measurement as expected from the probability of "one in four hundred". Any new measurement more than three standard deviations from the mean should ring alarm bells.

We can usefully quantify the origin of different values of standard deviation for Almen arc height determinations. In order to do that we use the term called "variance". Variance is simply σ^2, where σ is the standard deviation. The advantage of using variance is that total variability is simply the sum of the variances of the contributory factors. The total variability of repeated Almen arc height values, $\boldsymbol{\sigma^2_T}$, is made up of the separate variances due to strip variability, measurement errors, and variations in applied peening parameters. Hence, we have that:

$$\sigma^2_T = \sigma^2_S + \sigma^2_M + \sigma^2_{AP} \qquad (1)$$

where S, M and AP refer to strip, measurement and applied peening parameters respectively. Almen strips are produced to very close tolerances so that the σ^2_S contribution should normally be very small. "Premium grade" strips will produce a smaller variance than "standard grade" strips (other factors being equal). The σ^2_M contribution depends upon the quality of the Almen gage and the operator's skill/assiduousness. With good equipment and careful attention to detail, σ^2_M should also be relatively small. The major factor contributing to variability would then be predicted to be σ^2_{AP}. During actual shot peening there will always be some variation of the parameters that would affect strip deflection. Examples are: air pressure fluctuation, variations in flow rate and shot size (as when a batch of new shot is working its way through). Equation (1) quantifies contributions to total Almen strip measurement variability.

Consider, by way of illustration, two examples—A and B—reflecting good and poor combinations of factors respectively. Table 4 shows the results of applying equation (1) to hypothetical values (expressed in units of thousandths of an inch) of peened Almen strip.

Table 4. Effect of separate variances on total variability, σ^2_T, of Peened Almen strip deflection.

SET		σ^2_S	σ^2_M	σ^2_{AP}	σ^2_T
A (Good)	Variance	0.0001	0.0009	0.009	0.01
B (Poor)	Variance	0.0016	0.01	0.078	0.09

For the values given in Table 4, the applied peening variability predominates.

Data variability can, and should, be minimized by careful attention to all three contributory factors.

Bias

One obvious source of bias is the original strip curvature or "prebow". The origins and minimization of bias include: support ball wear, zero error and gage calibration over the full working range.

PEENING INTENSITY

Peening intensity is, perhaps, the most important statistic that we must deal with. It is estimated using a set of data comprised of four or more arc heights of Almen strips peened with nominally constant peening parameters. This procedure is, of course, familiar to all shot peeners. Fig.8 has the usual factors of a Solver Suite program with 99% confidence limits added. Each individual data point is subject to variability. Making repeat measurements at the same peening time would reveal the degree of variability. Careful attention to measurement factors can reduce, but not eliminate, the variability of each data point. The number of strips in the set is, however, important because it affects the variance of the derived intensity value. A larger number of data points in a set will improve the accuracy of the peening intensity estimate.

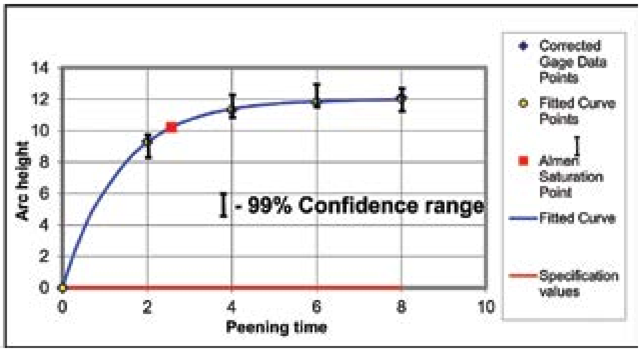

Fig.8. Variability of measured data points within a 99% confidence range.

CONCLUSION

Statistics is a subject that pervades everyday life. Several of the factors relevant to shot peening have been presented in this article. Consideration of those factors should feature in practical peening operations.