WHAT TO TEACH ABOUT CONTROL CHARTS AND PROCESS CAPABILITY IN BASIC STATISTICS

Bruce L. Bowerman and Richard T. O'Connell, Miami University
Bruce L. Bowerman, Dept. of Decision Sciences, Miami U., Oxford, OH 45056

In this paper we discuss what we should teach about control chart techniques and process capability in basic statistics. We also discuss how the topics might be taught in terms of a case study.

1. Introduction

The implementation of statistical process control involves achieving two related (but different) goals. The first goal is to achieve a state of "statistical control". A statistically controlled process is a process that displays a consistent amount of variability about a constant mean. In practical terms, the process basically operates in the same fashion over time. Statistical control is often demonstrated by using $\bar{X}$ and R chart techniques. Once achieved, this state of statistical control does not necessarily imply that the process is operating well enough to meet customer requirements. Control only tells us that no unusual process variations are being observed. The second goal is to achieve "capability". A capable process is one that is able to meet customer requirements (or product specifications). When both of these goals have been met, the process consistently meets customer requirements.

2. Using a Case Study to Teach Control Chart Techniques

The following is a modification of a case in Mendenhall, Reinmuth, and Beaver [3]. Their case was in turn based on a paper by James C. Seigel, "Managing with Statistical Models", SAE Technical Paper No. 820520, Society for Automotive Engineers, Inc., Warrendale, PA, 1982.

The Fort Motor Company produces the various parts for its vehicles in many different locations and brings these parts to central assembly locations. It is vital that these parts be within specification limits in order that they be assembled into a working, nondefective entity. Ford was having serious problems with a process used to harden the fuel pump eccentric of a 3.8-liter, V6 engine camshaft. The process produced inconsistent case hardness depth, causing 12% rework and 9% scrap. Excessive drill bit breakage occurred at the next operation in which an oil hole was to be drilled near the hardened fuel pump eccentric.

The hardening process was automated. However, an electric coil used in the process could be adjusted. To begin study of the process, one of Ford's problem solving teams selected thirty consecutive samples of n = 5 hardened fuel pump eccentrics. For each eccentric, the team measured the hardness depth at the eccentric lobe "nose". For each sample, the team calculated the mean $\bar{X}$ and range R of the n = 5 hardness depth readings in the sample. A plot of the thirty means and ranges obtained is given in Figure 1.

A process is said to be in statistical control if it displays a consistent amount of variability about a constant mean. Since the sample means and ranges in Figure 1 seem to exhibit substantial variability, the above process is probably not in statistical control. We will soon calculate control control limits to determine more precisely where the process has gone out of control (changed its variability and/or mean level). However, the problem solving team obviously did not have these limits to work with as it collected its samples because (as we will see) we use the samples to calculate the limits. Nevertheless, the team did take various actions to bring the process back into control when an intuitive assessment of the mean and range or other conditions indicated the process had gone out of control.

Specifically, at point A, which corresponds to a low average and high range, the power on the coal was increased from 8.2 to 9.2. At point B, which corresponds to an average and range that do not seem to indicate a problem, the team discovered and straightened a bent coal. At point C, which corresponds to a high average and high range, the power on the coil was reduced to 8.8. At point C, which corresponds to a low average and high range, the coil shorted out and needed to be straightened. At this point, the team devised a gauge to check the coil spacing at the camshaft. At point E, which corresponds to a low average, the team decreased the spacing between the camshaft and the coil. At point F, which corresponds to a low average and high range, the first coil was replaced with a second coil of the same type.

Note that most of the actions were taken when particularly large or small sample averages and/or ranges indicated that there were unusual sources of process variation. Such unusual sources are called <u>assignable causes</u> of variation. <u>Assignable causes</u> are intermittent or permanent changes that are <u>not common to all process observations</u>. These causes can often, as in the present case, be remedied by local supervision.

To more precisely determine when the process was out of control, we can calculate control limits. Suppose that a population has mean μ and standard deviation σ and is (or is close to being) normally distributed. Then, the probability is high that the mean of a sample of n observations randomly selected from this population will be in the interval

$$[\mu - 3\frac{\sigma}{\sqrt{n}}, \mu + 3\frac{\sigma}{\sqrt{n}}].$$

If a sample mean $\bar{X}$ falls outside of this interval, it is reasonable to conclude that the true population mean μ (and/or the population standard deviation σ) has changed. For reasons to be soon discussed, we estimate μ and σ by using the following procedure. We select K samples, each consisting of n process observations. Here, (in most situations) K should be at least 25 and n should be at least 3 or 4. Also, each sample should be a <u>rational subgroup</u> (a set of observations obtained while process conditions are not allowed to change substantially). Then if $\bar{X}_i$ and R_i denote the mean and range of sample i (for i = 1,2,...,K), we calculate

$$\bar{\bar{X}} = \frac{\sum_{i=1}^{K} \bar{X}_i}{K} \text{ and } \bar{R} = \frac{\sum_{i=1}^{K} R_i}{K}.$$

We estimate μ by $\bar{\bar{X}}$ and σ by

$$\hat{\sigma} = \frac{\bar{R}}{d_2}$$

where d_2 is given in Table 1 (see Burr [1]). It follows that the interval

$$[\mu - 3\frac{\sigma}{\sqrt{n}}, \mu + 3\frac{\sigma}{\sqrt{n}}]$$

is estimated by

$$[\bar{\bar{X}} - A_2\bar{R}, \bar{\bar{X}} + A_2\bar{R}] \text{ where } A_2 = \frac{3}{d_2\sqrt{n}}$$

(see Table 1).
The upper and lower control limits for the $\bar{X}$ chart are

$$UCL_{\bar{X}} = \bar{\bar{X}} + A_2\bar{R} \text{ and } LCL_{\bar{X}} = \bar{\bar{X}} - A_2\bar{R}$$

Without giving the logic, the upper and lower control limits for the R chart are

$$UCL_R = D_4\bar{R} \text{ and } LCL_R = D_3\bar{R}.$$

See Table 1 for values of D_4 and D_3.

For example, $\bar{\bar{X}} = 4.29$ and $\bar{R} = 1.80$ for the thirty consecutive samples of n = 5 hardened fuel pump eccentrics in Figure 1. Therefore, since Table 1 tells us that $A_2 = .577$, and $D_3 = 0$, and $D_4 = 2.115$, we have

$$UCL_{\bar{X}} = 5.35 \qquad LCL_{\bar{X}} = 3.25$$

$$UCL_R = 3.80 \qquad LCL_R = 0$$

These control limits are illustrated in Figure 2.

Examining Figure 2, we see that there are a substantial number of sample means and sample ranges outside of the control limits. One must be careful in interpreting the meaning of such results. To see this, we begin by considering a simpler situation. Suppose that we had observed the $\bar{X}$ and R charts in Figure 3. Before using the $\bar{X}$ chart, one should examine the R chart. Note that since the sample ranges are inside the control limits, the process is in control with respect to its variability. This means that the process is exhibiting a consistent amount of variability. A consistent amount of variability should exist before using the $\bar{X}$ chart control limits.

$$LCL_{\bar{X}} = \mu - 3\frac{\sigma}{\sqrt{n}} = \bar{\bar{X}} - 3\frac{(\bar{R}/d_2)}{\sqrt{n}}$$
$$= \bar{\bar{X}} - A_n\bar{R}$$

$$LCL_{\bar{X}} = \mu + 3\frac{\sigma}{\sqrt{n}} = \bar{\bar{X}} + 3\frac{(\bar{R}/d_2)}{\sqrt{n}}$$
$$= \bar{\bar{X}} + A_n\bar{R}$$

This is because these control limits assume that σ remains constant over time. At this point, it is important to point out why we estimate σ in the $\bar{X}$ chart control limits by using $\bar{R}/d_2$ rather than by using the sample standard deviation S obtained by lumping all of the observations in the K samples of n observations together into one sample. Note from Figure 3 that, because the process is out of control with respect to its mean, the range of the observations obtained by lumping the observations in sample 12 and 15 together into one sample is substantially larger than the average of the ranges of the observations in samples 12 and 15

$$\frac{R_{12} + R_{15}}{2}.$$

Therefore, if the process is out of control with respect to its mean, the standard deviation of the sample obtained by lumping together all of the observations in the different samples will provide a substantially larger estimate of the population standard deviation σ than $\bar{R}/d_2$, which employs the average of the ranges of the K samples of observations. Note that the number of observations in each of the K samples is fairly small (3,4, or 5) and observations are (usually) taken close together in time so that each of the samples will be a rational subgroup (a set of observations obtained while the process conditions are not allowed to change substantially). In this way we hope to make the ranges of the K samples fairly consistent and a reflective of the true process variation. Using the average of the ranges will hopefully give an accurate estimate of the standard deviation of a process that exhibits consistent variability. Of course, the ranges of the K samples might not be fairly constant. This what happened in Figure 2. Note that various sample ranges are outside of the control limits. Therefore, the process is out of control with respect to its variability. Specifically, since various R_i's are quite large

$$\bar{R} = \frac{\sum_{i=1}^{K} R_i}{K}$$

provides a larger estimate of the population range and a larger estimate $\bar{R}/d_2$ of the population standard deviation that would be obtained if the process were in control with respect to its variability. This means that the control limits for the $\bar{X}$ chart and for the R chart are more spread apart than they would be if the process were in control with respect to its variability. Therefore, we might not be detecting all of the sample means and ranges that indicate the process is out of control. Furthermore, a large $\bar{X}$ accompanied by a large R does not necessarily mean that the population mean level has changed. For example, consider $\bar{X}_{12}$ and R_{12} in Figure 2. Although $\bar{X}_{12}$ is larger than $UCL_{\bar{X}} = \bar{\bar{X}} + A_2\bar{R}$, the large R_{12} indicates that the process at the time sample 12 was observed was exhibiting more variability than the variability reflected by $A_2\bar{R}$ ($\bar{R}$ is, remember, the average of the ranges). Therefore, we do not know whether the fact that $\bar{X}_{12}$ is larger than $UCL_{\bar{X}}$ is due to the fact that the population mean level has increased or to the fact that the population variability has increased or to both facts. Another way to look at this is to note that a t-test would tell us to reject the null hypothesis that the population mean level at the time sample 12 was taken was constant at $\bar{\bar{X}}$ if

$$\frac{\bar{X}_{12} - \bar{\bar{X}}}{S_{12}/\sqrt{5}} > 3$$

that is, if

$$\bar{X}_{12} > \bar{\bar{X}} + 3(S_{12}/\sqrt{5}).$$

Because the t-test uses the standard deviation of sample 12, which is large because R_{12} is large, it would probably not allow us to reject the null hypothesis that the population mean level has changed. This does not mean that the t-test is a better procedure than control charts. In fact, the t-test potentially hides information that is provided by the control charts. In particular, if we use the t-test without noting the movement of the sample means and ranges over time, we might not note that a sample standard deviation (S_{12}) substantially larger than the standard deviations of other samples is what causes us to not reject the null hypothesis of constant mean level at the time sample 12 was taken. Therefore, we might be fooled into believing that the process is in control. In contrast, the $\bar{X}$ and R control charts make clear both a large $\bar{X}_{12}$ and a large R_{12}. Thus, we know that a problem does exist. In general, although it sometimes initially occurs that control limits such as the ones in Figure 2 are the best we can do, such limits provide substantial information about which sample means and ranges are out of control. This indicates what assignable causes need to be investigated and remedied. Indeed, we recall that at point A through F in Figure 2, Ford took actions to remedy the indicated problems. If we examine the actions taken at points A through E, we might say that the problem solving team has learned that the power on the coil should be roughly 8.8 and that it is important to monitor the spacing between the camshaft and the coil and to check for bent coils. This knowledge can hopefully be used to keep the process in control. If we believe that use of this knowledge will for the most part eliminate the assignable causes, we would recompute new $\bar{X}$ and R chart control limits by eliminating the $\bar{X}_i$'s and R_i's that were beyond the control limits and monitor future means and ranges by using the new control limits. However, recall that at point F the team replaced the original coil with a new coil of the same type. If it is felt that the new $\bar{X}$ and R chart control limits calculated from the original coil apply to the new coil, we can monitor the initial performance of the new coil by using these limits and then compute newer limits by using samples obtained from the new coil. When the problem solving team collected thirty samples of n = 5 observations, the results and control limits in Figure 4 were obtained.

Since no sample means or ranges are outside of the control limits, the process seems in control. It seems that the assignable cause variation has been removed. What remains is <u>usual process</u> variation. Usual process variation results from a combination of common causes of variation and chance causes of variation. <u>Common causes</u> of process variation are significant sources of variation that influence all process observations. A common cause of process variation might be obsolete equipment. For example, a poorly designed coil. <u>Chance causes</u> of process variability are small influences that cause variation even if we assume that all production conditions are being held as constant as humanly possible. Even though the process is in control (assignable causes have been removed), the usual process variation might cause the process to not meet <u>individual</u> product specifications. Once the process is in control, we may judge whether the process is meeting individual product specifications by calculating the <u>natural tolerance limits</u>.

$$[LNTL, UNTL] = [\bar{\bar{X}} - 3\frac{\bar{R}}{d_2}, \bar{\bar{X}} + 3\frac{\bar{R}}{d_2}].$$

Note that the above interval is an estimate of the interval $[\mu - 3\sigma, \mu + 3\sigma]$. Therefore, if the process is in statistical control, and if the population of individual process measurements is not far from being normally distributed with mean μ and standard deviation σ, the above interval contains substantially all <u>individual</u> process measurements (<u>almost all product specifications are written for individual measurements</u>). Thus, we wish to have both natural tolerance limits for a statistically controlled process inside the required product specifications. This tells us that substantially all individual process measurements are within the product specifications. As long as the process remains in control (as indicated by the $\bar{X}$ and R charts), we can be confident that we are meeting specifications. Using $\bar{\bar{X}} = 4.43$ and $\bar{R} = 1.6$ from Figure 4 and $d_2 = 2.326$ from Table 1, the natural tolerance intervals are

$$[LNTL, UNTL]$$
$$= [4.43 - 2.06, 4.43 + 2.06]$$
$$= [2.37, 6.49].$$

Specification limits for the hardness depths were 3.0 mm to 5.9 mm. The natural tolerance limits indicate, therefore, that the process does not produce almost all hardness depths within specification. Assuming the hardness depths are normally distributed with mean $\bar{\bar{X}} = 4.43$ and standard deviation

$$\frac{\bar{R}}{d_2} = \frac{1.6}{2.326} = .6879$$

it follows that

$$P(X < 3.0 \quad \text{or} \quad X > 5.9)$$

$$= P(Z < \frac{3.0-4.43}{.6879}) + P(Z > \frac{5.9-4.43}{.6879})$$

$$= P(Z < -2.08) + P(Z > 2.14)$$

$$= .0188 + .0162$$

$$= .035.$$

Therefore, the correct process is producing 3.5% out of specification hardness depths. At this point, the problem solving team had the coil redesigned to reduce the variability of the hardness depths. The $\bar{X}$ and R charts in Figure 5 were obtained.

The process is in control. The natural tolerance limits are [LNTL, UNTL] = [4.45 - 1.30, 4.45 + 1.30] = [3.15, 5.75]. Recalling the specification limits of 3.0 mm to 5.9 mm., the process is now <u>capable</u> of producing almost all individual hardness depths within specification.

3. A Summary of What to Teach

With the increasing use of the statistical techniques we have outlined above, it is imperative that we include a discussion of statistical control charts and process capability in basic statistics courses. Topic coverage should include all of the following: 1) Definition of common causes, chance causes, and assignable causes of process variability.
2) Construction of $\bar{X}$ and R charts.
3) Interpretation of $\bar{X}$ and R charts.
4) Discussion of the concepts of statistical control and process capability and how these concepts differ. 5) Calculation of natural tolerance limits for a process. 6) Interpretation of the natural tolerance limits and comparison of these limits with the specification limits. It can also be useful to teach the calculation and interpretation of the capability index C_{P_k}. See Mendenhall, Reinmuth, and Beaver [1989]. In addition, some instructors might wish to cover control charts for attribute data. These might include d-charts (defectives charts), p-charts (fraction defective charts) and c-charts (defects charts). Again see Burr (1976).

REFERENCES

1) Burr, Irving W., <u>Statistical Quality Control Methods</u>, Marcel Dekker, Inc., 1976.

2) Deming, W. Edwards, <u>Out of the Crisis</u>, Massachusetts Institute of Technology Center for Advanced Engineering Study, 1986.

3) Mendenhall, Reinmuth, and Beaver, <u>Statistics for Management and Economics</u>, Sixth Edition, 1989, PWS-KENT, Boston, MA.

Note: All control charts in this paper where copied (with permission) from Mendenhall, Reinmuth, and Beaver [3] (see pages 809 and 810).

Figure 1: Plot of the Thirty Means and Ranges

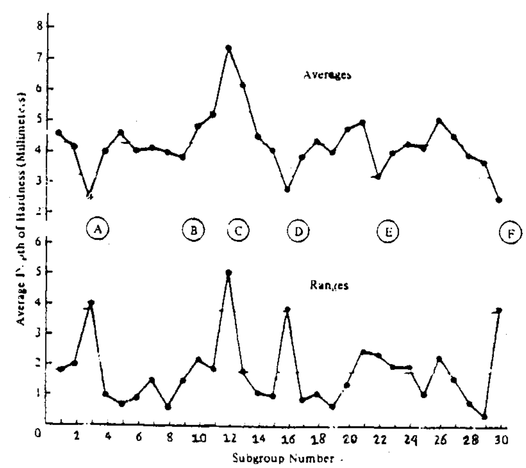


Table 1: Values of A_2, d_2, D_3, and D_4

n	A_2	d_2	D_3	D_4
2	1.880	1.128	0	3.267
3	1.023	1.693	0	2.575
4	0.729	2.059	0	2.282
5	0.577	2.326	0	2.115
6	0.483	2.534	0	2.004
7	0.419	2.704	0.076	1.924
8	0.373	2.847	0.136	1.864
9	0.337	2.970	0.184	1.816
10	0.308	3.078	0.223	1.777
11	0.285	3.173	0.256	1.744
12	0.266	3.258	0.284	1.716
13	0.249	3.336	0.308	1.692
14	0.235	3.407	0.329	1.671
15	0.223	3.472	0.348	1.652
16	0.212	3.532	0.364	1.636
17	0.203	3.588	0.379	1.621
18	0.194	3.640	0.392	1.608
19	0.187	3.689	0.404	1.596
20	0.180	3.735	0.414	1.586
21	0.173	3.778	0.425	1.575
22	0.167	3.819	0.434	1.566
23	0.162	3.858	0.443	1.557
24	0.157	3.895	0.452	1.548
25	0.153	3.931	0.459	1.541

Figure 2: Control Limits for the Sample Means and Ranges in Figure 1.

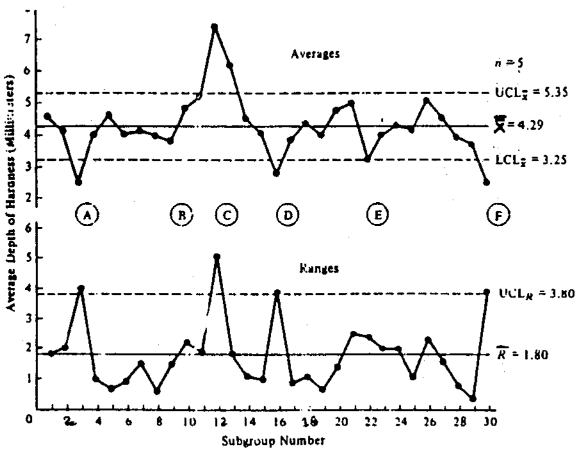


Figure 5: Control Charts Obtained by Using the Redesigned Control

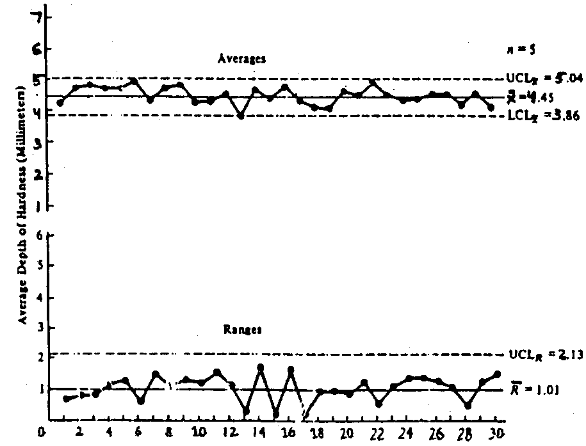
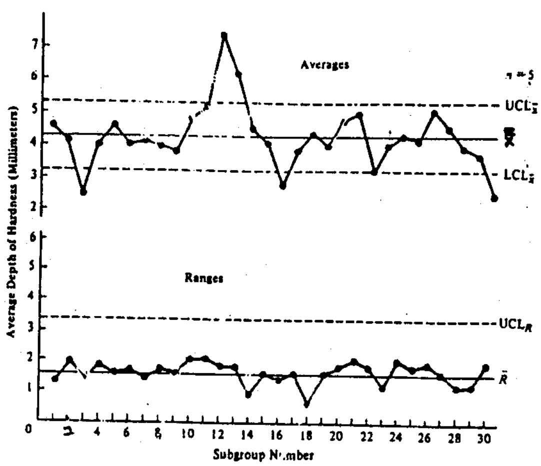


Figure 3: Control Charts Illustrating Ranges in Control and Means Out of Control



Range of sample 12

Range of Sample 15

Range of Samples 12 and 15 Lumped Together

Figure 4: Control Charts Obtained by New Coil of the Same Type as the Original