The Power of Histograms

Helping you understand the nature of the problem

by Sheila Julien, Senior Associate, Conway Management Company

All processes have some variation. We see it everywhere. A process may *typically* require 20 minutes, but will sometimes take 22 minutes, 18 minutes, or maybe even 35 minutes. Volume of calls, transactions, or orders may *average* 2500/day, but some days will see more volume and others less. A machine might be set to cut material into 2-foot widths, but actual widths may vary by 3/4 of an inch or so.

To understand how a process is functioning (or mal-functioning) we often need to understand the nature of the variation. How widely does the process vary? Within what range? Does the process variation follow a *normal* distribution — by that we mean, does the most frequently occurring outcome equal the calculated "average" outcome and is the outcome equally likely to be above average as below average?

Histograms are diagnostic tools helping us to better understand the nature of the variation.

They are often confused with Pareto charts, because both are displayed as bar charts. But while a Pareto chart will graph different categories (such as Product A, Product B, etc.) or causes (such as "lost orders," "missing customer number," etc.), histograms ***always*** have quantitative ranges (such as "1 to 10," "11 to 20," etc.) along the x-axis. Histograms are actually more similar to run charts because they both tell you about the range and average of individual outcomes in a time period. But histograms tell you the variation in aggregate and run charts tell you about the variation over time.

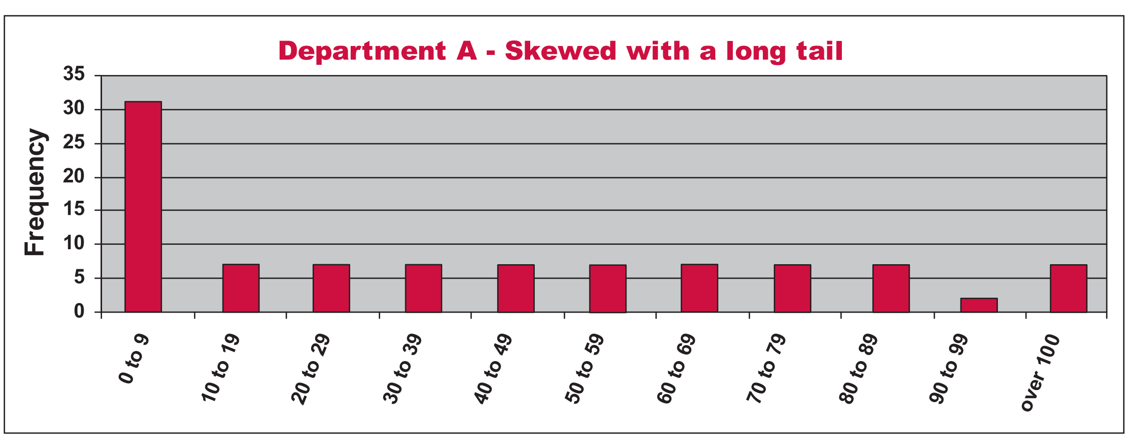
Why does the distribution of the process variation matter?

Here is an example:

Four different Customer Service Departments respond to telephone calls in 40 seconds, *on average*. Their surveys say that their customers are satisfied with response time 50 seconds or less. Since each department averages 40 seconds, there should be few complaints. But all four are receiving more and more complaints. Why? What should they do about it? Hire more Customer Service Reps? Maybe, but maybe not. In this example, the four organizations have four ***different*** underlying problems. The histograms help point them in the right direction.

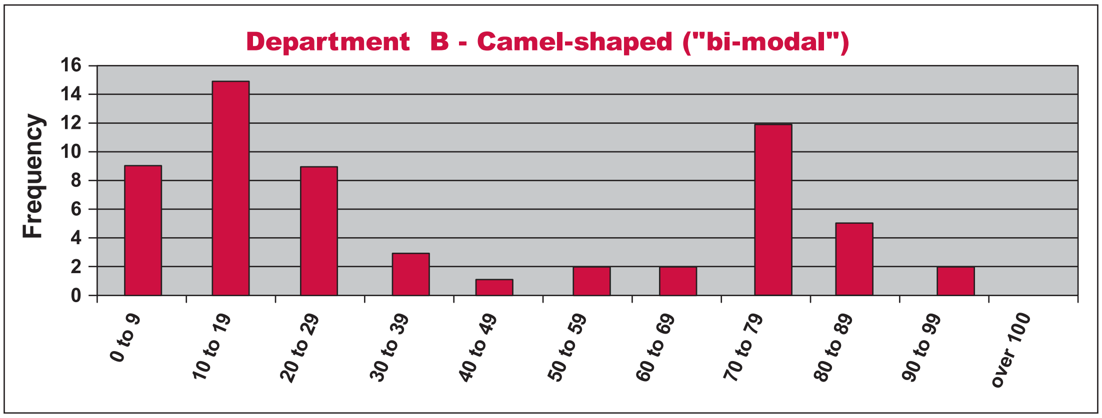
In **Department A**, as in all four cases, the average response time is 40 seconds - well within the target. But the histogram shows you that although the average response time is 40 seconds, the most frequent wait time is under 10 seconds and nearly 40% of calls have to wait more than 50 seconds. This is sometimes called a "cliff" distribution. When you see this distribution, you realize you should expect many complaints despite a satisfactory "average" performance. Whenever you have a distribution such as this, the calculated average has little relevance.

Furthermore, this particular distribution shows that for many of the calls, the department has more than enough people



available and other times they have too few. That is, resources are inefficiently scheduled and do not match call patterns. Perhaps reps all work the same 8 hour work schedule, rather than following call patterns. The department should test the idea that call patterns differ from scheduling practices. If this is found to be true, the department could significantly reduce complaints over wait time by realigning the work schedules to better match the call patterns.

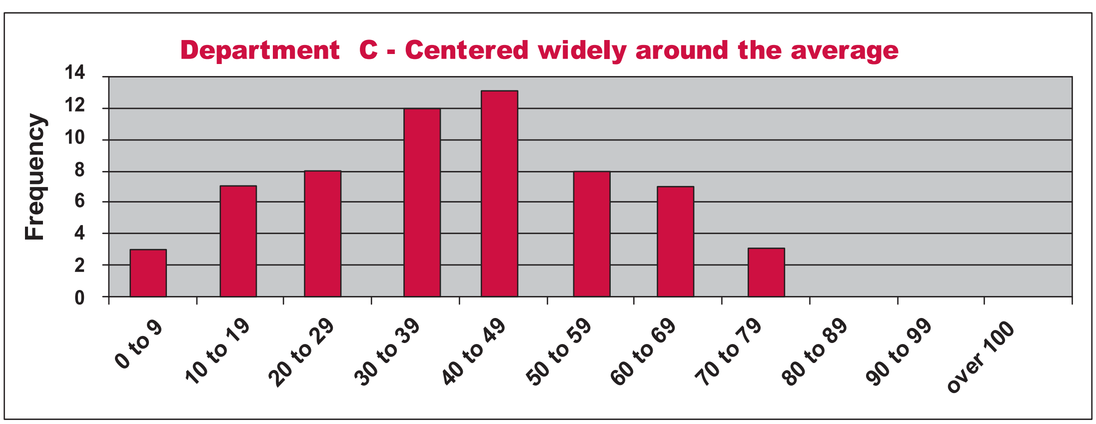
In **Department B**, the average response time is also 40 seconds. But the histogram for this department's call wait times shows that almost nobody waits 40 seconds. Either the call is answered pretty quickly (10-20 seconds) or the caller will probably have to wait 70 seconds or more. This "camel-shaped" distribution is sometimes called bi-modal, and if you have this sort of distribution, any calculation of an "average" is worse than meaningless - it is misleading. In this particular example, the calculated average is actually the least likely wait time.



When you have a bimodal distribution, there is a good chance that you have a mixture of two different types of events. It suggests that in one type of situation, wait times range between 0-30 seconds and in another, wait times range between 70 and 90 seconds. The next step would be to develop and test hypotheses about why there would be two different ranges of outcomes. Do they differ by shifts? Morning vs. afternoon; lunch hour vs. non-lunch hour? The department should ask the people closest to the work for their ideas about why wait times are either quite short or too long but seldom in between. Then segregate the data along those lines and test whether they indeed have two different processes — and then work on the subset that yields the unsatisfactory results.

In this example, Department B might segregate the data into calls that came in between 11:30 and 1:30, when the staff is going through lunch rotations — and discover that calls arriving between these hours have significantly longer wait times. It is clear that to reduce the number of complaints, the department need not add resources throughout the whole day, but only to find ways of increasing coverage between 11:30 and 1:30.

In **Department C**, the average response time is also 40 seconds. And unlike Departments A & B, most people who call Department C actually wait about 40 seconds. This process produces a "normal" bell-shaped distribution. Unfortunately, the bell is too wide. Even though the average and the most likely occurrence are below the customer requirement, the variation around the average is

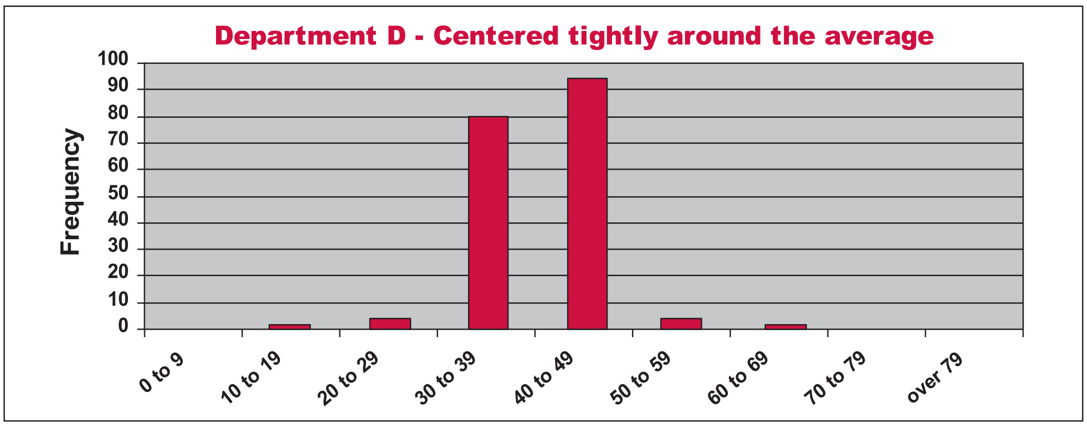


vidual calls frequently exceed the customer satisfaction threshold.

Department C could do two things to reduce the number of complaints about wait time. They can either make changes to shift the average wait time to 20-30 seconds or they can try to find and reduce the causes of variation — so that fewer call wait times are significantly higher than the average. Reducing the width of the variation is usually the most cost effective approach.

Department C will want to develop and test ideas about what contributes to the variation in call wait time. Does the experience or training of the telephone reps affect the wait time? Tools? Environment? Scheduling? Perhaps all of these and more affect this result. Find and correct the biggest contributor to variation in wait time to gradually reduce the frequency with which response time is much above the average.

In **Department D**, like Department C, the average wait time, 40 seconds, is also the most likely wait time. But call response time varies within a much narrower range for Department D. Instead of wait times averaging 40 seconds plus or minus 20 seconds, Department D wait times average 40 seconds plus or minus 5 seconds.



If Department D is receiving an increasing number of customer complaints despite the facts that the average (40 seconds) is 10 seconds better than what we think the customer requires and the variation around the average is very tight, then the best next step is to re-examine customer specifications.

Department D's problems appear to be due to changing customer expectations. Wait times that were satisfactory for callers last year may no longer be acceptable. Department D will not be able to reduce customer complaints about call wait times until they recalibrate their internal targets to the customers' new expectations and change the internal processes to shift the average downward.

For all four departments, the presenting problem was the same: More and more customers were complaining about the wait times when they called customer service — despite the fact that the average wait time was 40 seconds, well within the target that the most recent survey data said that customers would accept. But the underlying situation for each department is different. The histogram helps illustrate the nature of the underlying problem so that the departments can focus their resources on the right thing. ■